A contrastive study of English and Spanish synthetic diminutives

Anna Theresa Hägg

A thesis presented to the Department of Literature, Area Studies and European Languages

UNIVERSITY OF OSLO

May 2016

A contrastive study of English and Spanish synthetic diminutives

Anna Theresa Hägg

UNIVERSITY OF OSLO
May 2016

Supervisor: Kristin Bech

A contrastive study of English and Spanish synthetic diminutives

Anna Theresa Hägg

http://www.duo.uio.no/

Print: Reprosentralen, University of Oslo

For daddy,

who shared my enthusiasm for words

Abstract

This corpus-based study examines the expression of synthetic diminutives in the often assumed "non-emotive" language of English and the emotive language of Spanish. The topic of diminution is approached through a quantitative perspective, looking at two fundamental concepts of diminution, i.e. productivity of formation and distribution of denotative meanings/semantic denotations. The two corpora used in this study is the *Corpus of Historical American English* and *Corpus del Español*, which respectively provide evidence for the manifestations of English and Spanish synthetic diminutives in naturally produced language in the genres of Spoken, Fiction, News, Magazines and Academic texts. The study confirms that synthetic diminutive formation is more productive in Spanish, as is hypothesised, but in addition, it discovers that English shows signs of productiveness when comparing the relative frequencies of individual diminutive lexemes in English and Spanish. Also, this study argues that the diminutive type of specialised diminutives can be found in the English corpus data, a diminutive category which previously has not been acknowledged in English diminutive research, but which already constitutes an important function of the Spanish diminutive expression. Hence, the present study contributes to the lack of empirical research in the understudied field of contrastive word formation.

Keywords: English synthetic diminutives, Spanish synthetic diminutives, contrastive word formation, contrastive analysis, corpus-based, *Corpus del Español, Corpus of Historical American English*

Acknowledgements

In the year I have spent writing my master's thesis there are a number of people that I would especially like to thank, as I couldn't have done it without them.

First and foremost, I would like to thank my supervisor, Kristin Bech, who has inspired and helped me, by providing concrete advice for improvement. I will also thank her for initially believing in my suggestion for a topic, for the many hours she has spent reading my work, and, lastly, for sharing of the joy of diminutives with me.

I will also thank my fellow students on the eighth floor of Niels Treschow's Building, who have been a tremendous support morally throughout this year, for being able to discuss and laugh and making the whole prospect of a master's thesis less daunting. I must, in particular, place a special thanks to Marta, who has helped me with the intricate meanings of Spanish diminutives, as she has lived among locals for a longer period than I have.

I would also like to thank my mother for never failing to ask me how my thesis was advancing.

Tables and figures

Abbreviations

Pmw: Per million words

AmE: American English

SpALL: All dialects of Spanish, peninsular and Southern American

TC: *tertium comparationis*

LL: log likelihood

RQ: research question

Corpora:

COHA: The *Corpus of Historical American English*

CDE: *Corpus del Español*

***Dictionaries*:**

OED: *The Oxford English Dictionary*

DRAE: *Diccionario de la lengua española*

Table of Contents

1 Introduction

1.1 Background and aims

This study aims to explore the expression of synthetic diminution, variously termed diminutive suffixation (i.e. base + diminutive suffix), in the productive language of Spanish and the often claimed "unproductive" language of English. The kind of contrastive English-Spanish diminutive study that will be undertaken here has only to a small extent been attempted previously. The main aim of this study is to discover whether, in fact, English is as unproductive on diminutive suffixation as research suggests, when compared to a highly productive diminutive language such as Spanish. Often, diminutive research has included English only as a point of comparison to the diminutive language looked at in that study because it is the language of research, and thus is used as a point of comparison, but not really because they find the English diminutive expression interesting. This study is, I argue, one of very few studies that treat English as a productive diminutive language in itself.

In recent research Schneider (2003) and Rusek (2005) have attempted to write off claims about the unproductiveness of English synthetic diminution. There is still much to do at this point, if one wants to properly "whitewash" the ill reputation of English as an unproductive diminutive language. The focus of this study is to contrast English with Spanish diminution and thus reveal some facts about English diminution that previously have not been demonstrated.

Along the same lines as Schneider (2003), I argue in this study that the only viable way to claim anything about diminutives is through empirical research, i.e. through the study of naturally occurring language, either spoken or written, but definitely not constructed. As Schneider (2003: 69) confirms:

> As it is not produced for research purposes, material of this type is considered as material of the best quality and the most authentic, or even the only authentic, data type (...). Therefore, the claim could be made that, ideally, natural data should be analysed exclusively in research on diminutives (…)
>
> (Schneider 2003: 69)

To make up for the lack of empirically-based diminutive studies, and in addition, the lack of empirically-based contrastive word-formation studies (see chapter 3), and lastly, because I believe corpora to provide the most valid results in language research, the present study will apply corpora as the main source of data. As Sinclair (1987: xv) emphasises:

> Usage cannot be invented, it can only be recorded
>
> (Sinclair 1987: xv)

This study aims to shed light on the topic of synthetic diminution in English and Spanish by means of a contrastive approach employing corpus data as its main source of data.
The two diminutive aspects that will be investigated with regard to English and Spanish diminution are: (1) the productiveness of diminutive formation during the last two-hundred years, and (2) the repertoire of meanings that are expressed by diminutives. These diminutive aspects will mainly be studied through a quantitative approach. The two corpora that will be applied in the study are two large-sized comparable corpora, providing a quantitative base for the study, namely the *Corpus of Contemporary American English* (*COHA*) and *Corpus del Español* (*CDE*). The hypothesis that will be tested is the following:

> If it is the case that synthetic diminutives are more frequent in Spanish than in English, the present corpus findings from *COHA* and *CDE* should support this claim.

The hypothesis stated here will be tested by addressing two separate research questions, i.e. concerning each of the two diminutive variables stated, which together will be able to either reject or confirm the hypothesis. The first research question (RQ1) will look at diminutive frequencies and diminutive productivity and asks the following:

> Does English or Spanish demonstrate a higher relative frequency per diminutive lexeme?

The second research question (RQ2) will address the semantic distribution of diminutives, through the following question:

> In terms of semantic distribution, are the English or Spanish diminutives more innovative?

The analyses will be carried out in chapter 4.

1.2 Contrastive linguistics and practical applications

The present contrastive diminutive study belongs to the field of contrastive word-formation, which is a relatively new field of study and a merger of three different disciplines, i.e.

morphology, lexicography and contrastive linguistics. Exactly where the present study fits in is demonstrated in figure 1.1 here:

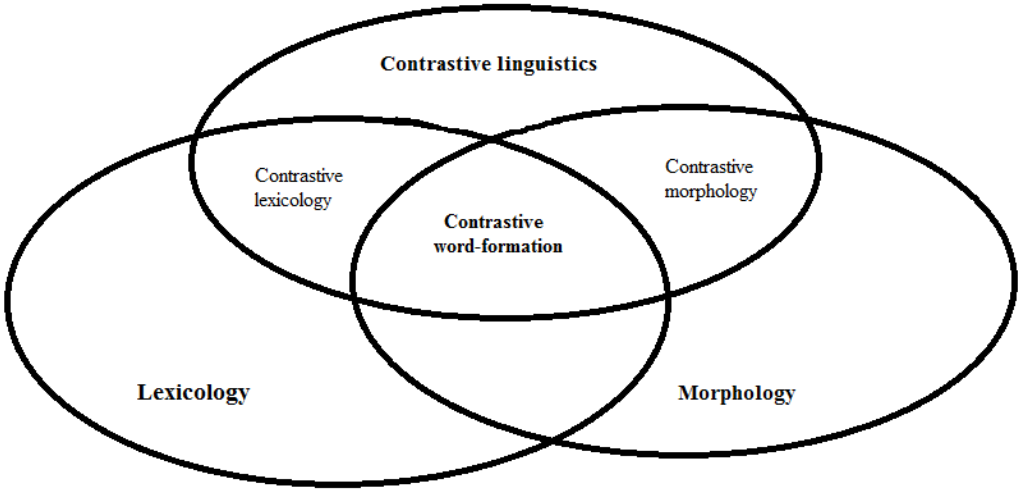

Figure 1.1: The three fields that have given birth to the field of contrastive word-formation

The current state of this contrastive word-formation is not ideal, to say the least, as the typical topics treated here have tended to avoid word formation topics altogether, focusing instead of purely grammatical and phonological aspects of language. The few topics within contrastive word formation that have been looked at, though to an insignificant extent, are suffixation, compounding and prefixation. Consequently, the present contrastive diminutive study will be a needed contribution, as it addresses one of these understudied areas of contrastive word formation. Additionally, contrastive word-formation lacks empirically based research, still remaining largely introspection-based. Similarly, from this perspective, the present study can contribute, as it in fact is one such empirically-based study.

1.3 Outline of thesis

The thesis is organised in four sections. Chapter 2 examines the topic of synthetic diminution, primarily focusing on English and Spanish. Chapter 3 accounts for the method used to carry out the analyses of this study, in addition to the type of data applied in the study, which is a corpus-based one. Chapter 4, which is the major focus of this thesis, concerns the two diminutive analyses of (1) lexeme frequency and productivity of formation in English and Spanish, and (2) semantic distributional patterns of English and

Spanish diminutives. Chapter 5 will sum up the findings of this study and raise some suggestions for further research.

2 Diminutive theory

2.1 Introduction

This chapter will introduce the topic of diminution in English and Spanish and acts as a foundation for the analysis of chapter 4. Fundamental concepts of diminutive formation will be discussed here, and issues concerning e.g. the "myth of English unproductiveness" will hopefully become a little bit clearer after this chapter. Throughout the discussion special emphasis will be given to synthetic diminutive formation, as this is the main concern of the present study. Other diminutive processes will only be mentioned.

This chapter is based on previous research of English diminutives, mainly carried out from an anglicist view, and consequently, the focus of this chapter is on English diminution. Spanish has been included as a point of comparison in most places. Other Indo-European languages will be mentioned as additional points of comparisons, but the emphasis is logically on English and Spanish. Unfortunately, much diminutive research is still introspection-based, which means that the present chapter may not portray English in a good light concerning productiveness. I hope that I have done a good job by highlighting the productive aspects of English diminution. If not, it will be evident why English is so often characterised as unproductive by most linguists. Ultimately, the two last sections will account for the diminutive repertoires of English and Spanish, which will later be applied in chapter 4 where the diminutive analyses takes place.

2.2 Diminutives in English and Spanish

Diminutive meaning can either be expressed synthetically, most commonly through suffixation (Sp. *¡pobrecito!* or *¡pobretín!)*, or analytically, through periphrastic constructions (Eng. *poor little thing!*). English is predominantly an analytical language, expressing diminution primarily through periphrastic constructions. Spanish, on the other hand, is a prototypical example of a language expressing diminution through suffixation, which is a general characteristic of all Latinate languages (Compare Sp. *besito,* It. *bacetto*, Por. *beijito,* to Eng. *little kiss*). In this respect, these two languages act as opposite poles in a continuum of diminutive expression. This chapter aims to shed light on the synthetic devices, i.e. diminutive suffixation, actually being used in English today, as many linguists claim this phenomenon to be poor, if not wholly non-existent, in the English language.

2.2.1 What are diminutives?

Diminutive formation is a near-universal concept attested across languages (Jurafsky 1996: 534; Haas 1972: 148; Schneider 2003: 2). In language, the diminutive category can serve many functions, both semantic and pragmatic ones (Dressler & Merlini Barbaresi 1994: 84: Augustyn & Gniecka 2011: 32). The present study will primarily focus of the semantic denotations expressed by diminution. According to the *Oxford English Dictionary* (*OED*), the semantic property of diminutives is to 'denot[e] something little' (2015). Cross-linguistic research confirms that the general function of diminutives is to express smallness (Jurafsky 1996: 534; Sifianou 1992: 157; Taylor 2003: 312). Diminutives are secondly known to express endearment and affection, often by suffixation as in hypocoristic forms and kinship terms, e.g. Eng. *Johnn-y, grann-y, dadd-y-kin-s* and Sp. *Juan-ita, abuel-ita, papa-íto* (Gooch 1967: 40). English also expresses diminution analytically through periphrastic constructions as in *dear little child*.

Diminutive formation is often said to originate from the domain of child language (e.g. Jurafsky 1996: 553, see 2.2.3). In the world of children, diminutives are applied as a means to signal affection, as expressed by e.g. *Petey, Gabrielito* (synthetic diminution). It is said that diminutives make the world less frightening by making it smaller and friendlier, and is thus used in conversations with children depicting the world as such (Sifianou 1992: 158). This use of diminutives can be seen in example (1), where as much as three different diminutive expressions is applied in the one and same sentence, i.e. suffixation, reduplication and analytic periphrastic construction:

(1) *Daddy* will only be away for a *teeny-weeny little* week, dear.

Diminutive use has thus spread from child language to other domains of language. Its resourceful nature makes (e.g. anything can be cute; small is relative) the diminutive category prone to serve different functions in language, though largely restricted to informal areas of language. The Spanish language has an extremely well-developed set of meanings included in its diminutive repertoire, and can be used to "to produce a favourable reaction in the person addressed" (Gooch 1967: 2), act as "mitigators", which reduce the amount or the effect of something, making the situation "emotionally manageable" as Gooch (1967: 3) terms it. Spanish diminutive can also mark social distance and politeness by their application in requests, offers and orders, as in *Un momentito, por favor* 'Just one moment, please'. Also, another diminutive function in Spanish is that of diminutives which have acquired a

specialised or changed meaning, as e.g. *mesa* 'table'< *meseta* 'tableland'. This type of diminutives is not expressive like most other diminutive types, but denotes rather factual meaning. The special use of diminutives says much about the extent to which diminutive suffixes, or for that matter, the role of derivational suffixation in Spanish.

English can also be said to be innovative as the language applies diminutives for pragmatic effects in the language (see e.g. Dressler & Merlini Barbaresi 1994). Augustyn and Gniecka (2011) look at such pragmatic functions of diminutives, and point out that diminutives are often applied to express irony:

> (2) We know you are going through grief and sorrow without dear *computie* working but hope you get through it soon Bastian.
>
> (Augustyn & Gniecka 2011: 31)

Such pragmatic meanings are best understood from the context, as the meaning of *computie* in isolation could have referred to, e.g. 'a cute computer' (the suffix *-ie/-y/-ey* typically expresses affection, see 2.2.6). Thus, as argued by Augustyn and Gniecka (2011), to account for the full range of possible meanings for diminutives, one has to include semantic, as well as pragmatic meanings (2011: 32). In addition, Schneider (2003) lists the following range of pragmatic functions of diminutives, i.e. "'playfulness', 'sympathy/empathy', 'understatement', 'euphemism', 'sarcasm', and 'stylistic choice'" (2003: 51). According to Schneider (2003: 51), the study of such pragmatic diminutive uses is most successfully accomplished through the study of corpus data, not dictionary listings, as has been the traditional approach to diminutives in research.

2.2.2 Nine diminutive strategies

The diminutive category as synonymous with diminution through suffixation stems from "traditional grammars originally used in description of Latin" (Schneider 2013: 137). To be able to treat the concept of diminution in world languages, not only in Latinate ones, a much broader definition is necessary. Diminution should include all linguistic forms denoting the semantic property of 'smallness' (Schneider 2003: 57), including phenomena such as syntactic modification, reduplication and compounding, all of which are formal means used to express diminution in language. This thesis acknowledges this distinction, but will, for the sake of simplicity, apply the term diminution more loosely.

Haas (1972) and Rosiak (2013) together name nine different diminutive expressions, which are the diminutive strategies attested so far in languages (148; 291), illustrated in the table below:

Table 2.1: Diminutive strategies attested across languages[1]

No.	Diminutive strategy	Examples	Languages
1	Derivational suffixation	Sp. *abuelita* 'granny', *tradecillo* 'rather late' (Lat. Am *tardecito*), Ger. *Hundchen* 'little dog', Eng. *cutie, kinglet*	Spanish, German, English, Italian, Portuguese, French, Polish, Russian, Greek, Creek, Natchez, Wiyot, etc.
2	Periphrastic constructions	Eng. *little boy*, Ger. *kleine Stadt* 'small city', Sp. *chica pequeña* 'small girl', Norw. *bitte liten gutt* 'tiny boy'	English, German, Spanish, Norwegian, Greek, Polish, etc.
3	Compounding	Ger. *Kleinstadt* 'small town', Wel. *afon* 'river' > *corafon* 'rivulet', Eng. *baby tree, dwarf tree*, Norw. *småby*	German, Welsh, English, Norwegian
4	Consonant/vowel symbolisation	Eng. *tiny* > *teeny*, Wishram *i-šgan* 'cedar board' > *wa-skan* 'box' > *wa-ck'un* 'cup', Yurok *pontet* 'ashes' > *pənčəč* 'dust',	English, Wishram, Yurok, Nez Perce, Wiyot
5	Grammatical displacement	Eng. Does *she* want mommy to wash *her* little hands?/Does *it* want mommy to wash *its* little hands*?*	English
6	Reduplication	Eng. *teeny-weeny, Annie-Pannie*, Snohomish[2] *ċaliʔ* 'heart' > *ċaċəʔi* 'little heart'	English, Snohomish, Nez Perce, Salish

[1] Examples taken from Robins 1958: 13 – 14, Haas 1972:148 – 151 and her own field notes (of the Natchez language), Rosiak 2013: 291.

[2] Snohomish is a Native American language, belonging to the Lushootseed languages.

7	Truncation[3]	Eng. *Johanna* > *Jo*	English
8	Inflectional suppletion	Sw. *ndege* 'bird' > *kidege* 'little bird'	Swahili
9	Suppletion	Natchez *han*-ho·*ʔis* 'to build normal-sized ones' → *han*-helu·*ʔis* 'to build little ones'	Natchez

What is it, then, that decides which diminutive expression comes to life in the different languages? Evidently, as research suggests, it is the structural make-up of a given language that largely determines the diminutive form appearing in the language. Some languages prefer morphological expressions, conveying a lot of information through a single morpheme, while other languages instead apply periphrastic constructions to express new meanings rather than forming new words. Is it the same way with diminution. It can either be expressed synthetically through different morphological strategies or analytically through syntactic modification. These two categories may again be classified according to the form the expressions take either within the synthetic or analytic type. Different classifications of the various diminutive expressions tend to include roughly the same repertoire of strategies (cf. Haas 1972; Rosiak 2013). Haas (1972) lists six, while Rosiak (2013) describes eight diminutive strategies, and they concur on syntactic modification being the only analytic example of diminution. Thus, according to these two accounts, eight morphological diminutive expressions are attested in languages, namely consonant/vowel symbolism, reduplication, derivational suffixation, inflectional suffixation, compounding, suppletion, grammatical displacement and truncation (Haas 1972: 148; Rosiak 2013: 291).

[3] This type is not always regarded as a diminutive form (cf. Schneider 2003 and Wierzbicka 1991?)

Diminutive suffixation – form and meaning

Derivational suffixation is, however, the method that is the most commonly associated with the diminutive category (Schneider 2003: 7) and is the one that has received the most attention in research so far.

Derivational suffixation is the most widespread synthetic method in Indo-European, with special prevalence in Mediterranean and Slavic languages, among them Spanish, whose structural make-up encourages this type of word-formation, which is being given special attention in this chapter. As the term relates, diminutive suffixation consists of the suffixation of a diminutive morpheme to the word base, thus adding a meaning of smallness to the original sense of the base word. Standard diminutive suffixation in Spanish has the following form: *casita* 'small house, (delightful) little house; cottage' (*casa* 'house' + *-ita* 'DIM.'), *pobretín* 'poor wee laddie (noun)' (*pobre* 'poor person' + *-ete* 'DIM' + *-ín* 'DIM') (Goosh 1967: 57; 78; *Collins Spanish Dictionary* 2005).

Diminutive suffixes are generally agreed to have two functions: one semantic and one pragmatic function (Dressler and Merlini Barabaresi 1994: 84; Jurafsky 1996: 534 – 535; Kryk-Kastovsky 2000: 165; Quirk *et al*. 1972: 994–995; Sifianou 1992: 157). The primary semantic function of diminutive suffixes is to express the idea of 'smallness'. This semantic meaning can also be applied metaphorically as markers of endearment or familiarity. Rosiak (2013: 289) explains that the idea of smallness both has literal and figurative interpretations, the latter sense being conveyed by metaphorical extension, while both meanings are realised by concrete, as well as abstract, forms, in the sense that 'to be small' is 'something dear, familiar' (Turner 2003: 172). In previous discussions of diminutives (cf. Kryk-Kastovsky 2000: 165), connotative meaning has often been confused with the pragmatic function of diminutives, e.g. represented by the English suffix *-y*/*-ie*/*-ey*, which has no inherent denotative meaning, but often indicates connotations of endearment, as in *cutie*, *doggy*. However, these two functions differ. Connotations are additional meanings, often expressing feelings and ideas invoked by the context, and they can be either appreciative or depreciative. This additional meaning is however, constant, and this is where pragmatic meaning differs; it is relative and changes according to the context, pre-existing knowledge, depending on the speaker's wish. Each utterance must, in fact, be interpreted "in relation to the speech situation" (Leech 1985: 13). Thus, the diminutive *bikey* in the sentence *my dear little bikey has stopped working* does not indicate endearment on the part of the speaker but instead

expresses irritation and contempt through the rhetoric tool of irony, normally known to express the opposite fact of what is actually stated.

Pragmatic effects of diminutives have normally been said to express the feature of 'non-seriousness', as first argued by Dressler and Merlini Barbaresi (1994: 144) in their extensive account of the morphopragmatic meaning of diminutives in German, Italian and English. The pragmatic meaning of 'non-seriousness' is communicated through the use of diminutive suffixes as a means to reduce one's bond and commitment towards the speech act and the force of the requests, warnings, etc., being performed (Dressler & Merlini Barbaresi 1994: 144). Also, the pragmatic property of 'non-seriousness' is related to the semantic quality of 'unimportant', often expressed by the English suffix *-let*, as in *lordlet*, *wifelet*, being a metaphorical extension of semantic 'small'. Ultimately, diminutives have the pragmatic effect of 'non-seriousness' when used to add playfulness and emotional flavour to the situation, thus downgrading the seriousness of the content, as often found in child-language (Dressler and Merlini Barbaresi 1994: 144).

Yet, the classification of diminutives as a morpho-pragmatic category is far from unanimous. Previous research (e.g. Lakoff 1987 and Taylor 2003) has sometimes failed to integrate semantics and pragmatics as interdisciplinary fields in their analysis of diminutives, turning a blind eye to the fact that meaning sometimes has to be understood through the contexts (Augustyn and Gniecka 2011: 38). Elements such as 'non-seriousness' and 'irony' can only be deduced from the context. However, fairly recently, in order to convey the specific nature of the diminutive category, a call for a diminutive analysis integrating both semantics and pragmatics was put into motion (Schneider 2003; Augustyn and Gniecka 2011). Diminutives frequently express several functions at once. Consequently, both the functional and formal aspects are crucial in the diminutive analysis and not each meaning in isolation (Schneider 2003), which requires a "crossing of boundaries" between semantics and pragmatics (Augustyn and Gniecka 2011: 47). The literal meaning of a word depends on the context, and thus in communication the word must be understood through the speaker and hearer play in order to interpret the literal meaning conveyed (cf. Langacker 1987).

Other diminutive strategies

This section will look at other types synthetic means to form diminutives, i.e. consonant/vowel symbolism, reduplication, grammatical displacement, compounding, truncation, inflectional suffixation and suppletion. These diminutive strategies will only be explained in brief as they constitute no large part of the present study. In addition, the only

analytic diminutive expression attested in language, i.e. periphrastic constructions will be discussed. This type is particularly common in English, but is not part of the focus of the present study.

The diminutive strategy of consonant/vowel symbolism has the function of increasing the diminutive meaning of an already diminutivised lexeme through replacing a consonant/vowel for another, as is illustrated here: *tiny* > *teeny*, where [ai] is substituted to [i]. A strategy which often co-occurs with vowel/consonant symbolism (therefore also treated in this paragraph) is reduplication. Typically in child language, the reduplication of first names first names occurs, i.e. *Annie-Pannie* and *Brinnie-Winnie* (Rosiak 2013: 291). Similarly, adjective constructions can receive consonant symbolism as in, e.g. *teeny-weeny* where [t] is replaced by [w] and the word is reduplicated (Haas 1972: 148).

Grammatical displacement comprises the act of substituting the second person pronoun for the third person pronoun, as in *Does she want mommy to wash her little hands?* instead of *Do you want mommy to wash your little hands?*. An even more intensified diminutive meaning by grammatical displacement can be achieved through changing the grammatical gender of the personal pronoun to the neuter, instead of the third person, often found when speaking to a toddler, as in *Does it want mommy to wash its little hands?* (Haas 1972: 149).

Compounding of a lexeme denoting smallness with a noun is used to express diminution in some languages. German makes extensive use of this type of diminution, as found in *Kleinstadt* 'small town' (Rosiak 2013: 291). English only has a few instances of diminutive compounding as found in e.g. *baby tree and dwarf tree (Bagasheva-Koleva 2010, online article)*. Typically, English resorts to syntactic modification in such contexts, applying the adjectives *small* and *little* to convey the meaning of smallness.

Truncation, also referred to as shortening or clipping, is another type of synthetic word-formation process used to form hypocoristics. As with hypocoristics in *-ie*, the truncated form is applied to signal familiarity or for the want of a more informal name of address, which is common both in English and in Spanish. It has the function of decreasing the social distance in the relation (Schneider 2003: 147). Especially in reference to adults, truncation is a frequent form of address, as opposed to the more common way of suffixing *-ie* to a name in reference to children to signal affection. Examples of truncated names, also called nicknames

or hypocoristics, are *Jo* > *Johanna*, *Chris* > *Christopher* or *Christine*, *Josu*[4] > *Jesús*, *Edu* > *Eduardo* (Prieto 1992: 144). In Peninsular Spanish dialects, this type of truncation is very productive, whereas Latin American dialects tend to use the last part of the full name as the form of address, as in *Chepa* (also *Pepa*, *Pita*) > *Josefa*, *Tína* > *Cristina* (Boyd-Bowman 1955: 357; 340).

Lastly, a few other marginal synthetic expressions are attested in language. Swahili can add diminutive meaning through inflectional suppletion as in *n-dege* 'bird' → *kidege* 'little bird'. Suppletion is the replacement of a word element, i.e. an auxiliary verb, by a word element with diminutive meaning, to add diminution. This use is not found in English.

The analytic diminutive alternative is expressed through syntactic means, i.e. periphrastic constructions. This process typically consists of the combination of an adjective expressing 'small' + a noun, which thus form a periphrastic construction. It is the adjective that marks the diminution and inflicts this feature on the noun (Schneider 2003: 122). In English, periphrastic constructions are often considered the only diminutive strategy, the adjective *little* extensively used to fill the function diminutive marker. Another neutral diminutive marker is *small*. In addition, the following adjectives sometimes occur in this function, the first five in informal contexts, while the remaining six are applied in formal situations: *tiny*, *wee*, *teensy*, *teeny-weeny*, *teensy-weensy* and *diminutive*, *minute*, *miniature*, *minimal*, *lilliput*, *petite* (Schneider 2003: 124).

2.2.3 Domains

Diminutives operate on many different language-specific areas. A common feature of all diminutives is that they most often appear in informal speech. Studies investigating the influence of culture on diminutive formation reveal how "expressive derivation interacts with speech acts", in languages which in general show a large amount of affection in speech, which again has an impact on the use of diminutives (Wierzbicka 1991). Languages rich on emotive suffixation apply diminutives to a range of different domains, i.e. in a sociolinguistic sense (Fishman 1964), while in less productive languages the diminutive suffixation occur in fewer domains (e.g. English).

[4] Examples of truncated names from the Spanish language are taken from Prieto (1992: 144).

Child language

The most common domain in which diminutive suffixation is applied cross-linguistically is in child language, variously termed "baby talk, motherese, (nursery) teacherese […] [and] child-centred speech" (Dressler & Merlini Barbaresi 1994: 173). According to Dressler and Merlini Barabaresi (1994: 173), in order to classify as child language, the child has to be either the speaker, addressee, side-participant or a non-present topical referent in a given situation, represented by the latter circumstance in (3):

> (3) Would the child ever let me live down the time the *little* dogg*ie* had that *teensy-weensy* seizure because I hadn't been wearing my glasses?
>
> (*COCA*, Fiction, 2012)

The use of diminutive suffixation in the domain of child language has been attested in research of various languages. This section will look at the use of synthetic diminutives in a few selected languages which serve illustrator examples of the occurrence of diminutives in this domain (some only in mentioning), i.e. English (Schneider 2003; Dressler & Merlini Barbaresi 1994; Haas 1972), Polish (Wierzbicka 1991), Greek (Sifianou 1992) and Spanish (Goosh 1967; Lang 1990).

In general, as briefly discussed in 2.2.1, diminutives occurring in this domain are applied as a means to signal affection to the child, presenting the world as a friendly place, and thus making it smaller. Lexical items particularly prone to accept diminutive endings are those that are directly or indirectly connected to young children, as extensively found in the diminutivisation of first names (Sifianou 1992: 158). The diminutivised first names are often termed hypocoristics (pet names), which are the diminutive forms occurring most regularly in English in this domain, as in *Petey*, *Annie*, etc. This diminutivised version of the first name often becomes the only term of reference when addressed by close friends, family or relatives.

In Polish, the formation of hypocoristics of first names is particularly productive. A first name can obtain a multitude of different versions, each diminutivised form expressing a marginally different attitude or mood (Wierzbicka 1991: 51). For instance, the proper name *Anna* has the following hypocoristic forms: *Ania, Anka, Aneczka, Anusia, Anu´ska, Anusie´nka, Anulka, Anuchna, Anusiatko.* A mother might address her daughter as *Ania* while a boyfriend would use the name *Anuchna*.

In addition, the pragmatic use of hypocoristic forms includes jocular and ironic uses. This use is sometimes overlooked in analyses in which semantic functions are the primary

concern. Lastly, animals are also commonly spoken of by endearing names, such as *little doggie* (Kryk-Kastovsky 2000: 171).

In Spanish, diminutive *-ito*/*a* is the most common suffix applied for affectionate meaning in child language, of the vast repertoire of diminutive suffixes existing in Spanish. The suffix is so extensively applied in this domain, that *-ito*/*a* is being avoided in more formal contexts, for fear of coming off as unprofessional or informal in tone. In addition, Lang (1990: 102) mentions two other Spanish suffixes, the allomorphic variants *-ica* and *-eta*, of *-ico* and *-ete* (for full accounts of the two suffixes, see 2.2.8). Formations in *-ica* and *-eta* are often used by children themselves, which Gooch (1967: 10; 70) terms "schoolboy language", normally expressing pejoration, as in *cobardica* 'cowardy-cat, funker' (from *cobarde), acusica* (from *acusar* '"tell-tale, sneak"), *llorica* '"cry-baby" (from *llorar*) (1967: 10; 70).

Food and drink

Another domain where synthetic diminutive formation occurs is talk about food and drink. Here, diminutive suffixes are used as a means to praise the food, its quality and related lexemes, in addition to making the customer feel that the amount of food or drink served or offered is small (cf. Wierzbicka 1991: 50), while also used for strategic means or gestures of politeness. Consider the sentence (4) below:

(4) ¿Vamos a tomar una *copita*?
[Let's go and have a glass of wine-DIM]
['Shall we go and have a (little) drink?']

(Gooch 1967: 58)

Here, the suffix *-ita* in *copita* 'nice/little glass of wine' can have various meanings. First, it can be applied to express affection and desire for drinking wine, as a personal opinion or as wine drinking is a cultural phenomenon. Secondly, it has the function of signalling that the invitation is not for a long night of drinking, but the though is to just take a glass or two. Thirdly, it can have the pragmatic effect of making the idea of going for a drink more inviting, sounding less of an effort.

Diminutives are often applied in these functions in the context of private meals, invitations to go for a drink or to eat, and lastly, in restaurant talk, where the occurrence of diminutives is particularly fruitful/frequent (Kryk-Kastovsky 2000: 169). We see here that the use of diminutives go beyond that of informal contexts, appearing also in the formal situation of restaurant talk.

English rarely applies diminutive suffixes in the function of praising food, but can apply periphrastic constructions to minimize the amount of food or drink, or as politeness markers, perhaps some affection, often in the form of the adjective *little*, as in (5):

(5) Come on let's have a *little* beer!

The diminutivisation of food and lexemes pertaining to this domain is especially common in e.g. Polish, Austrian-German and Spanish, in whose cultures food is regarded as important parts of the culture, also a cause of pride, which is thus reflected in the language through the application of diminutive suffixes, which, as previously discussed, have the ability to express affection.

2.2.4 Productivity

Measuring productivity of diminutive suffixes and diminutive strategies is a rather challenging task, especially as diminutives typically occur in informal types of language. Diminutive studies have typically not looked at informal language in research for various reasons (as briefly discussed in chapter 1). Much diminutive research is still based on outdated assumptions and impressions, which, in fact, is the case for most word-formation studies (Cannon 1987: ix). Diminutive studies based on lexicographical material, lack the width of language use. The fact is that dictionaries only accept lexicalised forms, including only the most common in conversational use (*sweetie*, *daddy*, *dearie*). Other types of formations are often left out of the dictioneries, as e.g. less common informal language (*mousie*, *birdy* 'small bird') and nonce formations (*punlet*, *egglet*, *wifelet* (Strang 1970: 90; Marchand 1969: 326)). It will be interesting to discover, which diminutive forms have been most commonly used during the last two hundred years, which the present study aims to find out through a corpora-based approach (see chapter 4).

It seems that in general, as argued by Schneider (2003), diminutives are "especially productive in those domains and discourse types which have been notoriously understudied to date" (82). Grammatical descriptions, e.g. Quirk *et al*.'s (1972: 993-994) *A Grammar of Contemporary English*, and other diminutive research have often reported just major generalisations about diminutive suffixation in English, providing examples from dictionaries, and not doing actual research. It is only recently that diminutive studies have started to include other domains of the English language, i.e. Augustyn and Gniecka's (2011) study of the irony expressed by diminutives, i.e. *computie*, where a corpus of English popular technical

forms was the material studied. Furthermore, Plag *et al*. (1999) state that the productivity of diminutives "can only be determined relative to particular domains and discourse types" (1999: 209) but not as a phenomenon in language.

The present study will apply the following definition of productivity, i.e. "A definition in terms of the possibility of forming new words […]" (Bauer 2001: 25), in the research of diminutives in chapter 4. In the corpus-based analysis taking place in chapter 4, a representative selection of suffixes will, among other things, be analysed for their ability to form diminutives in English and Spanish. Similarly, the productiveness of the diminutives in English compared to Spanish will be contrasted. As the present glance into diminutive theory shows, we cannot at once say that words ending in e.g. the suffixes *-ette* or *-ie* are diminutive formations, therefore notice must be taken to properly assess them for diminutive meaning. The suffix *-ette* can in addition to its status as a diminutive marker, denote the feminine gender as well, while similarly, the suffix *-ie/-y/-ey*, in addition to forming diminutives, also produces embellishing clippings (e.g. *rookie*). This study regards embellished clippings as non-diminutives (see discussion in 2.2.6).

Previous empirical research have listed the following diminutive suffixes to be productive, i.e. *-ette*, *-ey*, *-ie*, *-y*, *-let* and *-ling*, having produced diminutives such as e.g. *folderette*, *drilley*, *computie*, *celly*, *screwlet* and *loopling* (Augustyn and Gniecka (2011: 34). Augustyn and Gniecka (2011) look at diminutives from a corpus of English technical terms, which found these diminutives as used informally as applied in e.g. the computer language of internet discussion forums.

It is easier, however, to distinguish productive languages from one another. Here, taxonomies of suffixal repertoires can be compared, the flexibility that suffixes have to appear singly and serially, and the ease with which a suffix attaches to various grammatical classes. On all three accounts, Spanish is more productive than English. For instance, English morphology makes it impossible to add a diminutive suffix to some names, instead English compensate with other means, i.e. periphrastic constructions of names together with a kinship term such as *sweetie, honey, sweet pea, pumpkin, baby* (Kacmárová 2010: 19). Here, Spanish applies a suffix, for the address form *Anita* (for *Ana*), which in English would be *Anna pumpkin/sweetie*, etc.

2.2.5 Grammatical aspects

The standard grammatical form of diminutive suffixation is word base + derivational suffix (any inflectional suffixes are added to the right of derivational suffix(es)), as in i.e. *pig-let-s*. The base lexeme *pig* and the derivational morpheme *-let* express the lexical meaning, while the inflexional morpheme *-s* expresses additional grammatical information, here plurality. However, different diminutive patterns across languages exist which extend or explore the possibilities of the diminutive nature. This section will examine two grammatical variables of diminutive formation which differ across languages, namely multiple diminutivisation and rules for grammatical base of attachment.

Multiple diminutivisation

Diminutive suffixation can produce derivatives with one suffix, which is the standard diminutive expression (Eng. *booklet*), but diminutives can also be formed by several consecutive suffixes added to the base, variously termed multiple diminutivisation or recursive diminution. This phenomenon occurs extensively in Spanish (Sp. *chiquitito*, *pobretín*), but is far less frequent in English (Eng. *popsy*, *fatso*). Multiple diminution can occur in the form of the same suffix applied twice, as in Sp. *chiqu-it-ito* 'tiny, miniature', called reiterate suffixation, which is found in Spanish, but not in English. Secondly, two or more different suffixes can attach to the base, as in Sp. *carr-et-illa*[5] 'trolley, cart', Eng. *pregg-er-s* 'pregnant', and is very popularly used in Spanish (*bomb-ill-ita*) (Lang 1990: 93). Multiple diminution reinforces the semantic meaning of smallness or endearment of the word (cf. *chiquito* and *chiquitito*) and can also add additional connotative meanings (Kryk-Kastovsky 2000: 168). For instance, the addition of diminutive *-ín* to *pobrete* 'poor thing', enforces the already positive meaning *pobretín* has the positive 'poor little fellow' in a positive way, while *pobretón*, the augmentative suffix *-on* producing a negative meaning 'poor man/woman'

Languages which have a rich emotive morphology are known to be productive on multiple diminutivisation of words. This is particularly frequent in Romance languages, as

[5] The two diminutive suffixes *-eta* and *-illa* have been added to Spanish *carro* 'car' producing a new lexeme, *carretilla* 'trolley, cart', whose meaning is a metaphorical extension of 'car' (http://is.muni.cz/th/109655/ff_b/BAKALARKA.doc).

noted by Schneider (2003: 117), and also in Slavic languages. For instance, Polish demonstrates the use of double diminutives, sometimes bringing about a semantic change, e.g. from *cake* [Pol. *ciast-o*] to *cookie* [*ciast-ecz-ko*]. German multiple diminution can occur in the form of *-lein* as in *Maus-i-lein* and *Schatz-i-lein* (Kryk-Kastovsky 2000: 168; 172).

English has a less developed emotive morphology and therefore quite naturally makes less use of multiple suffixes in its expression of the diminutive. The common patterns of suffixal combination often occur in the form of *-ie*, *-s* and *-o* (Quirk *et al*. 1985: 1584). Poynton (1989) states that first names are a group that is particularly prone to multiple diminution in English, claiming that they "display the most elaborate set of morphological resources for forming diminutive forms anywhere in English […] ". In addition, Schneider (2003: 117) reports that recursive diminution occurs with common nouns as well:

(6) -er + -s → shampers
-le +-s → cuddles
-ie + -poo + -s → kissipoos

Australian English is the variety that produces the most productive patterns of multiple diminutivisation in English. Some derivations demonstrate occurrences of triple, even quadruple, diminutive formation, as in *Frang-le-kin-s* and *Mike-y-poo-dle-s*, which are not just nonce formations (cf. e.g. Gramley/Pätzold 1992: 289).

Choice of primitive base

Diminutive formation across languages prefers nominal bases as primitives. English diminutive suffixes are predominantly denominal; examples of suffixes attaching to nouns are *-ette* (*kitchen* > *kitchenette*), *-let* (*stream* > *streamlet*), *-ling* (*duck* > *duckling*) (Bauer 1983: 221). However, some diminutive suffixes demonstrate some irregularity, i.e. flexibility, in being able to attach to adjectival bases, as represented by the suffixes *-y*/*-ie*, *-ling* and *-o* as in *fatty*, *cutie*, *wierdo* and *weakling* (Dressler & Merlini Barbaresi 1994: 112). The suffixes *-ie*/*-y*/*-ey*, *-ette* and *-ling* also infrequently attach to verbal bases, as in *cookie*, *launderette* and *hireling*, a fact that Dressler and Merlini Barbaresi's (1994: 112) overlook in their account of English diminutives. In the case of *-ling*, the derivatives of adjectives and verbs usually have derogatory connotations (Quirk et al. 1972: 994; Huddleston and Pullum 2002: 1678).

Spanish diminutive suffixation, on the other hand, is much freer, allowing diminutive formations in the grammatical categories of nouns, adjectives, adverbs, pronouns and verbs. Some diminutive suffixes, i.e. *-(c)it*, can be added to all these five grammatical categories

(Steriopolo 2015: 15). In comparison, German and Russian diminutive suffixation is normally restricted to nouns (Steriopolo 2015: 18), as in English, while Polish diminutive suffixes can practically attach to any concrete noun, adding an element of affection (Szymanek 1989: 112), as in i.e. *masel-ko* 'butter-DIM' (Kryk-Kastovsky 2000: 168).

2.2.6 Repertoire of diminutive suffixes in English

Representative selections or generalisations of English suffixation reveal that only a small percentage of English suffixes have a diminutive meaning (see Quirk *et al*. 1972: 993-1008; Katamba 2005: 59-63). Grammarians and linguists often characterise synthetic diminution in English as being poor, and point out that the suffixes only constitute a minuscule part of the English derivational morphology (Quirk *et al*. 1972: 994; Dressler and Merlini Barbaresi 1994: 84).

The suffixes used to express diminution in English (today and historically), will here be accounted for according to their attested productivity in the language to date, ranging from high to low. Altogether, the inventory of diminutive suffixes in English consists of *-ie, -ette, -let, -ling, -kin, -een, -en, -*(*e*)*rel, -et, -o, -a, -s, -er, -poo* and *-pegs*.

-ie/-y/-ey

Research and grammatical descriptions of the suffix *-y*/*-ie,* (also *-ey*) describe it as a marker of familiarity or endearment, its main function being to form hypocoristics such as *Annie, Johnny, Susie*, etc., while secondly it is applied in child language (see 2.2.3) (Dressler &Merlini Barbaresi 1994: 112; Charleston 1960: 123; Huddleston & Pullum 2002: 1677; Dossena 2012: 3). In addition, Dressler and Merlini Barbaresi (1994) distinguish a third type where it frequently occurs, namely in kinship terms, such as *mummy, daddy, granny*, etc. (112). A forth, rather controversial use of the suffix *-y/-ie*/*-ey* applied to abbreviated forms occurring in all English varieties, will be discussed below in terms of its use in Australian English where this formation is exceptionally productive (Sussex 2004: 8).

The suffix appears in three variants, *-y, -ie* and *-ey*. The two former are more frequent in use, while *-ey* appears to a lesser extent. In hypocoristic forms the spelling *-ie,* from Scottish, is preferred, as in *mousie, dearie*. Proper names sometimes prefer one ending to the other, the female names *Sally* and *Anni*e being more common than the variants *Sallie* and *Anny* (*OED* 2015).

The *-y/-ie/-ey* suffix is generally agreed to have no proper semantic meaning;[6] instead linguists emphasise its role as a marker of emotional closeness, which is characteristic of hypocoristics and kinship terms such as *Jenny*, *dearie* and *granny* (Quirk *et al*. 1972: 995; Huddleston & Pullum 2002: 1677). In addition, the suffix can express irony or sarcasm, often in contexts where there is an element of irritation involved (cf. Augustyn & Gniecka 2011: 32-33; Dossena 2012: 4). In expressing these various functions, it differs from the general nature of English diminutives, whose primary purpose is to denote smallness and which very rarely express additional pragmatic functions (e.g. *-ette* 'small, compact'; 'imitation' and *-let* 'small, unimportant').

The grammatical characterisation of *-y/-ie/-ey* is as a denominal suffix, as it primarily produces nouns from nominal bases. It can occasionally also attach to adjectival bases, as in *brownie, goodie, softie* (e.g. 'a week-minded or silly person'), and marginally to verbal bases as in *cookie,* always producing nouns (Huddleston & Pullum 2002: 1677). However, it must not be confused with the deadjetival *-y* (type 2) 'like, full of, etc.', whose chief function is to transforms concrete mass nouns to gradable adjectives, such as *creamy*, *hairy*, *silky*, etc. (Quirk *et al*. 1972: 1002). The application of the diminutive suffix requires polysyllabic lexemes to be reduced to monosyllabic bases in order to produce a diminutive derivative as exemplified by *granny* for 'grandmother', *tatties* for 'potatoes' (Dressler & Merlini Barbaresi 1994: 112).

In terms of linguistic productivity, linguists ascribe this suffix to be the most productive diminutive suffix in English (Charleston 1960: 123; Dressler & Merlini Barbaresi 1994: 112; Huddleston & Pullum 2002: 1677; Schneider 2003: 87). In some statistics it even emerges as the most frequent suffix of all derivational suffixes in use in present-day English (Cannon 1987: 185). Sinclair (1991: 188) deems its high productivity rate to possibly be a result of its ability to combine with almost all names and nouns, which no other English suffix can (1991: 188).

Australian English and the use of suffix -ie/-y/-ey

A rather controversial use of the suffix *-ie/-y/-ey* is the application of the suffix to form abbreviations, producing embellished clippings of longer words, common in all English varieties (*movies* for 'moving pictures'). Australian English demonstrates a great increase in its use of abbreviations and suffixes to form diminutives and hypocoristics compared to other

[6] A few exceptions exist in the lexicalisations of *piggy* and *pinkie* 'little finger'.

languages (Sussex 2004: 8). One of the most common suffixes used in this function is *-ie/-y*/*-ey*, as seen in *firie* for 'fire officer', *lippie* for 'lipstick' and *brolly* for 'umbrella'. Recent creations such as *druggie* for 'drug-addict' and *greenie* for 'i.e. the curtain of a theatre'[7] also belong to this category, but are not particular Australian creations (Quirk *et al*. 1972: 995; Huddleston and Pullum 2002: 1677).

The controversial issue at stake here is whether these Australian abbreviations, which very much look like diminutives, really express smallness or endearment, which is the core meaning commonly associated with diminutives. Wierzbicka (1991) argues that the function of the Australian abbreviations of *-ie(s)/-y*/*-ey(s)* is solidarity markers, not expressing diminution or endearment (1991: 55). She elaborates that the overall meaning of these forms expresses good-humour, mirroring cultural aspects of Australian philosophy, such as the importance of "mateship", a desire for informality, light-hearted sarcasm and distaste for lengthy words (Wierzbicka 1985: 169; 1991: 56). Ultimately, Wierzbicka (1985) states that Australian abbreviations also differ in form, because the suffix is added to the truncated form and not the full form of the word as the "true" diminutive does (1985: 169). Dressler and Merlini Barbaresi (1994) and Sussex (2004) agree with Wierzbicka (1985; 1991), while Schneider (2003) argues that these Australian abbreviations behave in a diminutive manner. Some studies mention truncation as a means to form diminutives, most often forming hypocoristics, as in *Joseph* > *Joe*. The form *Joey* is indeed an embellished abbreviation of a longer word which behaves like a true diminutive, expressing endearment. Examples of this type favour Schneider's (2003) view of including these embellished clippings as proper diminutive formation.

Standard Scottish English

Standard Scottish English also demonstrates a real appetite for applying *-ie* to form new words, much more so than English, but it cannot be compared to the productivity of Australian English. The Scots language developed from Middle English in a separate direction, thus becoming a distinct, but similar Germanic language, of English. Common hypocoristics in SSE include various variations of the sort *lassie* (lad), *dearie*, *fishie*, *dishie*, *beastie*, etc. often occurring in child-talk and kinship terms for both humans and animals (Dossena 2012: 3). Illustratively, in a children's poem by Bulloch (1921/1970), almost every noun is diminutivised:

[7] The OED (2015) lists six meanings of *greenie* still in use.

Dance to your *daddy*, my bonnie *laddie*.
Dance to your *daddy*, my bonnie lamb!
And ye'll get a *fishie* in a little *dishie*
Ye'll get a *fishie* when the boat comes hame.

It is from Scottish influence that English has received the forms *laddie*, *dearie*, etc.

-ette

The diminutive suffix *-ette* is of French origin, adopted by English in the Middle English period, then also in the form of *-et*, a suffix which today has ceased to produce new lexis. Diminutive *-ette* (*-et*) arrives as a borrowing from Old French from the 1100 to the early 1800s, leaving a considerable mark on the English vocabulary (Charleston 1960: 121; *OED* 2015). The spelling kept its original French form, but the pronunciation is monosyllabic /-ˈet/.

Three semantic meanings can be distinguished from the suffix *-ette*. The first sense indicates small size, functioning as a true diminutive, as in *kitchenette, novelette* and *sermonette*¸ *colonette*. The second sense denotes 'imitation' especially of material, and can be considered diminutive derivatives in which the meaning of 'smallness' is extended by metaphor, imitation being "a little like the original thing", as in *leatherette, flannelette*. The third bulk of words that *-ette* appears in, denotes the feminine gender, as in words like *majorette, suffragette* and *usherette,* and must be distinguished from lexemes denoting sense one or two, actually expressing diminutive meaning (Quirk *et al*. 1972: 994; Bauer 1983: 119; Huddleston & Pullum 2002: 1677). The feminine marker *-ette* came into English in the 1850s (Schneider 2003: 92).

Today, the suffix *-ette* is still to a certain degree productive, in diminutive and emotional function, but most often as a feminine marker. It is particularly productive in the domains of advertisement and trade, exemplified by Nieuwenhius' (1985) collection of new lexis in *-ette*, such as *featurette* 'a short feature film', *snackette* 'a snack bar' and *drawbackette* 'a small drawback' (Charleston 1960: 121). However, its productiveness cannot be compared to that of the endearing suffix *-y*/*-ie*/*-ey*, which is by far the most productive suffix in use today. Interestingly, however, *-ette* is the only non-native diminutive suffix currently adding new lexis to the English vocabulary today (Leisi 1969: 89).

The chief grammatical function of *-ette* is that of a denominal noun suffix as seen in *kitchenette, statuette*, etc., though some derivations of *-ette* have verbal bases. *Launderette* is one such irregular derivation, its primitive base being the verb *launder* (Marchand 1960: 290; Schneider 2003: 93). *Dinette* and *slumberette* are sometimes regarded as derivations of the

truncated noun forms *dining-room*, **slumber-room* (unattested form) (Marchand 1960: 290), while Schneider (2002) suggests that *dinette* and *slumberette* could well have been formed on the verbal bases *dine* and *slumber* (93). Furthermore, most derivatives of *-ette* have disyllabic bases, the suffix seemingly refusing to attach to monosyllabic bases (Dressler & Merlini Barbaresi 1994: 113).

-let

The suffix *-let* is often considered a Germanic suffix (Schneider 2003: 97; Rotzoll 1910: 109), but is by some linguists said to be of French origin (Charleston 1960: 121). It functions as a denominal noun suffix, producing countable nouns such as *flatlet, islet* and *leaflet,* also often the young of animals, as in *piglet, froglet* (Quirk *et al*. 1972: 994; Huddleston & Pullum 2002: 1677; Katamba 2005: 62).

In terms of diminutive status, *-let* is often characterised as the "cleanest" and most significant diminutive suffix in English, because it almost exclusively is used to indicate diminution and its diminutive meaning is strongly felt. In comparison, other diminutive suffixes (*-ie* and *-ette*) are often ambiguous in their nature, expressing a range of different meanings, some of which can only be said to be vaguely diminutive, if at all, and thus compromising their diminutive status (Schneider 2003: 96, cf. Rotzoll 1910: 30).

The suffix *-let* can express two different meanings. It can indicate small size, as in the countable nouns *starlet*, *booklet*, the young of animals, as in *troutlet*, *froglet*, or a small variety, such as *droplet*, *leaflet*, etc. Secondly, it denotes the property of 'unimportant', often in a derogatory sense, such as *kinglet, princelet, lordlet*, i.e. in the meaning of a 'a petty prince' (Charleston 1960: 122; Quirk *et al*. 1972: 994; Dressler & Merlini Barbaresi 1994: 113). Some derivations of *-let* are considered to be of unknown origin, as they are not used diminutively but instead denoting the quality of 'resembling…' as *wristlet, anklet*, etc. (Huddleston & Pullum, 2002: 1677 – 1678).

In terms of productivity, it is variously described as a "moderately" or "marginally" productive suffix today, producing few new lexical items in present-day English (Dressler & Merlini Barbaresi 1994: 113; Huddleston & Pullum 2002: 1677). However, the frequency of *-let*-derivations still in use today is high, comparable to that of *-ette*, which is largely due to a particularly productive period in the latter half of the 19th century, where most of the derivations in use today, such as *booklet* and *leaflet*, were created (Schneider 2003: 96). In addition, it characteristically appears in the formation of nonce words, i.e. words which exist

only in restricted periods of time, such as *punlet*, *egglet* and *wifelet* (Strang 1970: 90; Marchand 1969: 326; *OED* 2016).

-ling

The Old Germanic suffix *-ling* usually denotes the young of animals, as found in *duckling*, *codling* and *gosling* similar to the use of the diminutive *-let*. The suffix *-ling* acquired this particular diminutive meaning in the Middle English period, possibly through Old Norse (ON) contact, which used the suffix extensively in this function. It is still used in Norwegian, as in words like *mannsling* 'poor, pathetic man'. Additionally, it can denote the young of plants, as in *sapling*, *seedling*. Furthermore, it often conveys contempt, primarily in reference to adult persons, as in *princeling*, *squireling*, *hireling*, which came into use in this meaning in the 16th century. This use is possibly also adopted from the ON, which used to speak of young animals with affection, and must since have developed to convey derogatory senses as well (*OED* 2015; Huddleston & Pullum 2002: 1678; Charleston 1960: 122).

Derivations only expressing small size are hardly used with *-ling*, existing only in nonce creations from the 19th century, such as *philosophling*, *thinkling* and *metaphysicling*, according to the *OED* (2015). In Old English (OE), however, it was not a diminutive suffix, but used to denote a 'person or thing of a specific kind or origin' or 'son of', a use which today no longer is productive (*OED* 2015).

The main grammatical property of *-ling* is to produce nouns from nouns, either in the form of animals or humans, as there are no inanimate derivations of *-ling* (Katamba 2005: 62). Other grammatical categories used to produce diminutives are adjectives and verbs, as the formations *weakling* and *hireling* illustrate, which typically express contempt. The suffix here demonstrates greater flexibility than other English diminutive suffixes, as it is able to form diminutives from other categories than nouns (Quirk et al. 1972: 994; Huddleston & Pullum 2002: 1678).

-kin

The suffix *-kin* is a Dutch borrowing, cognate to German *-chen*, coming into English in the 1300s (Jespersen 1942: 462). The first examples of this suffix appear in addresses of personal names, chiefly male, such as *Janekin*, *Malekin*, *Watekin* and *Wilekin* (c. 1250), which stayed popular for about a century, then went out of use. The suffix has both a diminutive and

endearing force and can express both affectionate and derogatory connotations, as implied from the context (Schneider 2003: 105).

These address forms have instead survived as surnames, often in combination with *-s* or *-son*, as in *Jenkins, Watkins, Wilkinson, Dickens, Dickinson*, etc. Only sporadic examples exist of the suffix in use with common nouns, as found in the historic forms of *baudekin, fauntekin* and *feudekin*. However, about a century ago the derivation of nouns with *-kin* was quite popular, as confirmed by Bradley (1908) stating that "we can, at least in jocular speech, add *-kin* to almost any noun to form a diminutive" (1908: 138).

Today, the suffix only has a limited use, which seems to be restricted to nicknames, the suffix attaching to disyllabic bases often ending in *-ie*, and is almost consequently followed by the suffix *-s*, as in *Katiekins, Lizzikins* (Schneider 2003: 105 – 106). Charleston (1960) and Schneider (2003) conclude that the only diminutive in *-kin* still having proper currency is the fossilised form *lambkin* usually applied only in the domain of child language (1960: 123; 2003: 105). Other words such as *jerkin, bumpkin, pipkin, gaskin, griskin, ciderkin*, only express slight notions of affection and their origins are unknown. Most of these nouns, and surnames listed above, are no longer morphologically analysable, and thus, in most of the cases the diminutive strength is lost altogether (Huddleston and Pullum 2002: 1678).

-een, -en, -(e)rel, -et

This group of suffixes can loosely be termed historical diminutives, as they used to express diminution in the English language, but have now lost their diminutive force. The majority of the remaining words containing *-een*, *-en*, *-(e)rel* or *-et* are fossilised word forms that can no longer be separated morphologically, as explained for the suffix *-kin*, as well, which is illustrated here by *colleen, kitten, doggerel, scoundrel, bullet* and *bumpkin* (Huddleston and Pullum 2002: 1678). A few exceptions exist in *maiden, cockerel, baronet, lambkin* and *manikin*, where the base and suffix have independent lexical meanings (Huddleston and Pullum 2002: 1678).

The Anglo-Irish suffix *-een* still exists in names and words of particular Irish descent, such as *colleen, mavoureen, pot(h)een, Squireen, Kathleen, Eileen, Maureen* and *smithereens*¸ where the diminutive meaning is still felt (Charleston 1960: 123). In addition, some examples of *-een* express 'imitation', as in *velveteen* and *sateen* (Huddleston and Pullum 2002: 1678).

-o, -a, -er, -s and irregular diminutives

The suffixes *-o, -a, -er* and *-s* are common in use with so-called embellished clippings, but their diminutive status is controversial, as discussed earlier in the case of the suffix *-ie/-y/-ey* (see Wierzbicka 1985: 169; 1991). Characteristically, the suffix *-o* occurs in words expressing the same function as *-ie*/*-y*/*-ey*, and competes with these more frequent diminutive forms. It is only, however, applied to describe males, whether pejorative adjectives or in hypocoristics of proper nouns, as in *fatso*, *wierdo* and *Stevio*. Contrastingly, formations of adjectives in *-ie*-derivatives tend to have positive implications, as in *kiddie, weirdie,* whereas diminutives in *-o* normally express pejoration, as in *kiddo, weirdo* (Schneider 2003: 111).

There is no doubt that derivations of nouns with the suffix *-o* producing hypocoristics and kinship terms are true diminutives, such as *Stevio* and *kiddo*. The controversy arises with *-o* being attached to clippings, identical to the *-ie/-y/-ey*-suffix currently popular in Australia (see previous discussion of the suffix *-ie/-y/-ey* in this section). Quirk *et al*. (1985: 1584) term these abbreviations familiarity markers, while Wierzbicka (1984) argues for their expression of solidarity. This use of *-ie* and *-o*, where the suffix attaches to abbreviated word forms, creating so-called embellishing clippings, as in e.g. *ammo* (from 'aggravation'), *arvo* (from 'afternoon'), I would argue, similarly like Wierzbicka (1984) and Quirk *et al*. (1985: 1584) that they are not true diminutives.

Other similar uses occur with the suffixes *-a*, -er, *-s* and a few irregular ones. They will not be included in this discussion.

2.2.7 Classification of derivational suffixes in Spanish

The primary way to produce words in Spanish is by derivational suffixation. It is common to distinguish between so-called emotive and non-emotive suffixes in Spanish (Lang 1990: 91; 123). Emotives, also termed appreciatives, constitute a rather small part of the roughly 200 derivational suffixes existing in Spanish. They are, on the other hand, extremely productive and have a special function in the language. It is to this group that diminutives belong. Emotive or appreciative suffixes are added to words to express affection towards an object or a person (abstract entities are not compatible with these suffixes) but do not alter the grammatical class of the base, which non-emotive suffixes frequently do. Some emotives can, however, change the gender of primitive base (*Nueva gramática de la lengua española* 2010: 163; Lang 1990: 91).

Traditionally, emotive suffixes are grouped according to their main function, being either diminutive, augmentative or pejorative in manner. The classification of emotive

suffixes into the three main categories is, however, a difficult task. Spanish emotives have the ability to express all three meanings at the same time, namely diminution, augmentation and pejoration. Consequently, the task of categorising them is a matter of interpretation, and what is more, the meaning of the emotive often changes, often according to the situation, or the common occurrence of semantic change. Goosh (1967: 5) argues that because diminutive meaning is hardly ever clear-cut, one should avoid assigning one meaning to each suffix, and instead view semantic classifications only as guidance, analysing the individual diminutive from the context it appears in.

2.2.8 Repertoire of diminutive suffixes in Spanish

The Spanish grammatical reference work, *La Nueva gramática de la lengua española* (2010), lists the following nine suffixes to be the most frequent diminutive suffixes in peninsular Spanish, *-ito/a, -illo/a, -ico/a, -uelo/a, -ín/ina, -ino/a, -iño/a, -ajo/a, -ejo/a* and *-ijo/a*. Gooch (1967) and Lang's (1990) accounts of the diminutive category name seven and six suffixes, respectively. The following five diminutive suffixes are found in all three grammatical accounts and will therefore be applied as the basic repertoire of diminutive suffixes in Spanish:

-ito/*a, -illo*/*a, -ico*/*a, -ín*/*ina*[8] and *-ete*

In general, only a small percentage of diminutive derivatives are included in the Spanish dictionaries, because new diminutives are created continuously at the same time as others go out of use. Thus, definitions of Spanish diminutives in either Spanish or English are hard to come by. Goosh's (1967) account of diminutive, augmentative and pejorative suffixes in modern Spanish is the main source of the examples and definitions used in the current description of Spanish diminutives. Here, the primitive bases are only given if the derivative has changed severely. His examples have been retrieved from modern Spanish literature, newspapers and magazines, and actual conversation, and are unmistakably an extremely useful guide to diminutive use and meaning in Spanish, especially for the non-native speaker (Goosh 1967: xi).

[8] The feminine variant *-ina*, of masculine *-ín*, has a homograph in the diminutive suffix *-ina*, from *-ino*, and must not be confused in meaning. Both are diminutive, but with slightly different uses.

The two main qualities diminutives express are smallness and affection, but Spanish diminutives have developed several more meanings, the phenomenon being so fruitful in the language, in contrast to, e.g. English. Spanish diminutives frequently express intensification (¡*es chiquitín!*), courtesy (¿*alguna cosita más, señorita*?) and adding of colour (*el mundillo del cine* 'the (narrow little) world of the cinema') (Goosh 1967: 7). The number of words needed to adequately render the same meaning into English demonstrate the wealth of meanings Spanish diminutives can express (Goosh 1967: x). Another characteristic of Spanish diminutives is that each suffix has its own distinctive shade of meaning, i.e. some being preferred in certain regions, some being more pejorative than others, and therefore the selection of a particular suffix in use for a certain context should be a careful one (Lang 1990: 102*; Nueva gramática de la lengua española* 2010: 166).

-ito/a

The most frequent diminutive suffix in the Spanish language is *-ito/a*. It has a neutral diminutive meaning, is the least regionally marked, and occurs in all Spanish-speaking communities across the world (Lang 1990: 103; Gooch 1867: 5; *Nueva gramática de la lengua española* 2010: 166). Primarily, it attaches to nouns, *bombita* 'small bomb, bomblet' often to proper nouns (hypocoristics), but also frequently to adjectives, in addition to some adverbs and participles (Lang 1990: 103: Gooch 1967: 110). It characteristically expresses smallness and endearment, more frequently expressing the latter as in hypocoristic *abuelita* 'granny', and the kinship term *amiguito* 'fine friend', in adjectives *acabadito* 'nice and finished', adverbs *bajito* 'very softly' and participles *los niños iban corriendito* 'the children ran off happily' (Gooch 1967: 37-65).

The suffix normally attaches to lexemes ending in *-a* and *-o* (*Nueva gramática de la lengua española* 2009: 638; 643). There are two allomorphic variants of *-ito*/*a* in *-cito*/*a* and *-ecito*/*a*, and the suffix also has some rarer forms in *-ítar, -itos* and *-citos* (*azuquítar* 'a little sugar', *Carlitos* 'dear/little Carlos', *Dolorcita*s). Some linguists claim these to be cases of infixation, where the morpheme *-it-* is inserted inside the base *Carlos* (*Carl -os + -it- > Carlitos*). The *Nueva gramática de la lengua española* (2009) treat it as a less common allomorphic variant of *-ito*/*a* (2009: 166).

It is, furthermore, the most frequent diminutive suffix found in the domain of child language (see 2.2.3 for a definition) as the popular stories of *Caperucita Roja* and *Blancanieves y los siete enanitos* illustrates (*Little Red Ridinghood* and *Snowwhite and the*

Seven Dwarfs) (Gooch 1967: 2). Its association to child language, does, however, often lead people to select another diminutive suffix or mode of expression in normal speech, choosing *bosque pequeño* or *bosquecillo* for 'small wood' (Gooch 1967: 9).

Some pragmatic functions of *-ito/a* include an ironic use, *problemita* from (*el*) *problema*. Feminine derivatives normally carry subtle undertones of irony, as in *frasecita* 'quite a sentence or phrase'. Masculine derivatives, on the other hand, often have stronger augmentative-pejorative connotations, as can be seen from *corazoncito* 'tiny heart', *qué añito* 'what a (terrible) year this has been' (Gooch 1973: 37-38; 58; 65).

The suffix can also cause a specialised or change in meaning, in both masculine and feminine lexemes. Examples of derivations with a specialised meaning are *tortita* 'pancake; waffle; flap-jack' from *torta* 'type of bun, cake or pastry' and *carrito* 'trolley, e.g. for carrying food to table' from *carro* 'cart, wagon'. Instances of derivatives where the meaning has completely been changed are *palomita(s)* 'popcorn' from *paloma* 'dove' and *llanito* 'Gibraltarian, native of Gibraltar' from *llano* 'flat; plain' (Gooch (1967: 46; 63-64).

In the attachment to adjectives in past participle function *-ito* expresses diminution and affection as seen in *ancianito* 'quite old, getting on rather in years', *delgadito* 'rather thin, slender' (Gooch 1967: 47-48). Derivatives of adverbs *-ito/a* normally have an affectionate meaning or have an ironic or augmentative-pejorative force, as in *bajito* 'very softly, very quietly', *despacito* 'very slowly, very gently, nice and slowly', (*hasta*) *lueguito* '(see you) in just a while, (see you) very soon'.

Finally, the young of animals also often receive the suffix *-ito/a* to express affection as in, describing animals with affection, as in *gatito* '(sweet) little cat', *burrito* 'small monkey' and *pajarito* 'small bird, (sweet) little bird'. This use of *-ito* denoting the small of animals corresponds to the use of the English suffixes *-let* and *-ling* as found in *piglet*, *froglet* and *duckling*, *gosling*, respectively.

-ico/a

Contrastingly, diminutive *-ico/a* is a regionally marked suffix, frequently occurring in the regions of Aragón, eastern Andalucía and in many parts of Central America. The suffix used to have a much more prominent role, especially in literature of older Spanish, where it enjoyed particular popularity, as can be seen from the following passage from *Celestina* (1499) (Lang 1990: 108; Gooch 1967: 25):

(7) -¿Qué dirás a esto, Párameno? *¡Neciuelo, loquito, angélico, perlica, simplecico! ¿Lobitos* en tal *gestico*? Llégate acá, *putico…*

(Espasa-Calpe 1955: 95)

A corpus search of the randomly selected lexical synonyms *angelito* and *angelico* confirms that is *angelito* is the most frequent of the two during the 1900s, with a frequency of 50 per mill. words, compared to 12 for *angelico* (*CDE* 2015).

Furthermore, *-ico*/*a* is only productive in its attachment to nouns, as in *besico* 'gentle/ nice/sweet little kiss', *pajarico* 'cute/sweet little bird', *nenica* 'cute/sweet little girl/girlie; darling, sweetie'. It does only marginally attach to the bases of adjectives *solico* 'all by oneself', verbs *llorica* 'cry-baby' and adverbs *tantico* '(just) a wee bit' (Gooch 1967: 65-70; Lang 1990: 108).

In terms of meaning, *-ico*/*a* express diminution, being slightly affectionate, and is close to the meaning of the standard diminutive *-ito*/*a* in use (Lang 1990: 108). In addition to this meaning, two other semantic uses occur. An intensifying use, common for Spanish diminutives, is particularly frequent with masculine adjectives, such as *malico* 'poorly', *pobrecico* 'poor sweet little…'. The allomorphic variant *-ica* is the only one used pejoratively, as in *miedica* (from *miedo*) 'chicken, coward', *cobardica* 'cowardy-cat, funker', which can be used as adjectives as well, and are modern creations popularly used by schoolchildren (Gooch 1967: 67; 70-71; Lang 1990: 108).

-illo/a

Particularly in earlier times, in ancient Roman times (c.0) and in the Golden Age (c. 1492-1660), *-illo* was the preferred diminutive variant (*Nueva gramática de la lengua española* 2010: 166; Lang 1990: 102). Today, it is still widely used, both in Spain and in Latin America, but it is often connected with Andalucía where it perhaps is more frequent than even *-ito* (Land 1990: 103).

The suffix *-illo*/*a* is fundamentally diminutive, but demonstrates a clear preference for light pejorative meaning, as in *regalillo* 'a modest/wretched little gift', *gentecilla* 'petty riff-raff', *escocesilla* 'wee Scots lassie'. Secondly, it is exceptional in its ability to produce lexemes with a specialised or changed meaning, as in *bolcito* 'small handbag', *cabecilla* 'ring-leader, rebel leader, partisan leader', hipocritlla 'bit of a hypocrite' (Gooch 1967: 8;

122-123). It does, however, occasionally occur in a pure diminutive sense, often affectionate, as in *blanquillo* 'small bench', *arenillas* 'small/fine grains of sand' (Gooch 1967: 5-6; 104).

In productivity, *-ito/a* most often attaches to nouns (*hombrecillo* '(insignificant) little man', *acentillo* 'slight/unpleasant little accent'), frequently also to adjectives (*alegrillo* 'quite cheerful, lively, gay', *fuertecilla* ' strongish; pretty strong; heavyish'), in a lesser extent to present participles (*creciendillo* 'gradually growing', *lloviznandillo* 'gentle drizzling', (*de*) *puntillas* 'on tip-toe') and rarely to adverbs (*tardecillo* '(pretty) latish') (Gooch 1967: 82; 97; 99; 103; Náñez Fernández 1973: 111).

-ín/-ina

The diminutive *-ín/ina* is a regional marker for Asturias and is similar to *-ito* and *-ico* in that it essentially expresses favourable, often affectionate meanings, as in *besín* 'gente/sweet little kiss' (cf. *besito* and *besico*). Pejorative uses, being less pejorative than *-ete* and *-uelo*, only appear in lexicalisations with specialised or changed meanings with the masculine variant *-ín* (Gooch 1967: 10; 72).

Its formation to different word classes is restricted in comparison to its rival diminutive suffixes. The only productive word class for *-ín/-ina* is nouns, producing some hypocoristics, too (*Jaime* > *Jaimín*). Attachment to other word classes is restricted, occurring only with the most common adjectives (*pequeñín* 'tiny, wee, tichy; tiny chap', *tontín* 'silly boy') and some verbs (*bailar* > *bailarín* 'very fond of dancing', *cantar* > *cantarín* 'very fond of singing; murmuring e.g. brook; sing-song e.g. voice') (Lang 1990: 107; Gooch 1967: 78 – 80). These deverbal derivatives are special in that they are formed on the full infinitive, not the base (*bail-*) as they normally would, and secondly because a change of word class occurs, from verb to adjective, which is uncharacteristic for emotive suffixes, as stated in the introductory section of 2.2.8 (Lang 1990: 107).

The most characteristic function of *-ín/-ina* is to form lexicalisations with specialized or changed meanings, as its rivals *-illo -ete* and *-uelo* also frequently do, as in *faldellín* 'short skirt; kilt' from falda 'skirt', *folletín* 'penny dreadfun; melodrama; cheap serial' from *folleto* 'pamphlet'. It is only in this function that the feminine allomorphic variant *-ina* occurs, exemplified by *madrina* 'godmother, sponsor', *sobaquina* 'under-arm sweat' from *sobaco* 'arm-pit' (Gooch 1967: 75-81).

Another distinctive quality of this suffix is its tendency to occur in combination with other diminutive or emotive suffixes, though only the masculine variant *-ín*, appearing last in

order, as in *pobretín* (*pobre* + *-ete* + *-ín*), *poquitín* (*poco* + *-ito* + *-ín*), *manojín* (*mano* + *-ajo* + *-ín*) (Lang 1990: 108). As previously mentioned in section 2.2.5, multiple suffixation reinforces or intensifies the diminutive meaning of the word.

-ete/a

The suffix *-ete*/*a* (allomorphic variant *-cete*, as in *meloncete*) is characteristically used to express jocularity and irony, as in *cerdete* 'jolly old pig, fine young pig', *pobrete* 'miserable pauper, poor devil' (Gooch 1967: 10; 141-142, *DRAE* 2016). It is considerably pejorative in meaning, having the strongest pejorative effect of the five Spanish suffixes discussed here, though the suffix *-uelo*, which is not discussed here, transcends *-ete*/*a* in "negativeness" (Lang 1990: 102).

Spanish only has one suffix in the form of *-ete*/*a*, but that does not mean that all lexemes ending in *-ete*/*a* are diminutives. Actually, quite a large number of lexemes ending in *-ete*/*a* in Spanish are direct borrowings from French. In French, *-et* is a noun- and adjective forming suffix, often producing diminutives. In addition, it can also be just an integral part of the stem, and a suffix, as in *ribete* 'border, addition, adornment' (*Collins Spanish Dictionary* 2005).

One of the main functions of Spanish diminutives is to reduce the effect or quality of words that e.g. are considered inconvenient or uncomfortable, or to diminish the quality of the denoted, as in the adjective *rojito* 'un poco rojo'. In derivatives of *-ete*, it is the seriousness of the primitive meaning that is reduced, adding a jocular tone or quality, often through irony, as in e.g. *amiguete* 'mate, buddy', *golfete* '(proper) little urchin/mischievous youngster', *mentirosete* 'lying bitch?' (*Nueva gramática de la lengua española* 2010: 169; Gooch 1967: 141). Many derivatives of *-ete*/*a* have, however, lost their diminutive meaning altogether, as e.g. *peseta* (from *peso* 'coin'), *juguete* 'toy' (from *juego* 'game').

The suffix allows for diminutive formation to a number of different word classes, though only regularly producing diminutives from nominal bases, often slightly jocular, as in *librete* 'jolly old book' *isleta* 'small island, islet', *placeta* 'small square'. Adjectival bases are relatively frequent, as in *guapete* 'quite handsome; handsomeish' (cf. *guapillo*), *alegrete* 'a bit gay or merry, quite or rather merry or high (from drink); pretty high'. Formation of verbal bases exists, such as *clavetear* from *clavar*, but are regarded a special type of diminutive suffixation, because the suffix *-ete* here occurs as an infix. Other grammatical classes are incompatible with *-ete/a*. (Gooch 1967: 141-153; Lang 1990: 105).

In addition, as the research in chapter 4 also will show, the suffix *-ete/a* frequently produces diminutives with specialised meaning. It is particularly the feminine variant *-eta* that is applied in this function. Some examples from Gooch (1967) are *arete* 'ear-ring' (from *aro* 'ring'), *colorete* 'rouge' (from *color* 'colour'), *tableta* 'tablet, bar' (from table 'board, plank), *coleta* 'pig-tail' (from *cola* 'tail'), *libreta* 'note-book' (from *tabla* 'board, plank') (1967: 141-153).

2.3 **Summary**

It has been shown in this chapter that diminutives, commonly represented by synthetic diminutives in research, have different manifestations in different languages. Spanish has been shown to demonstrate very productive patterns of formation, e.g. in being able to apply recursive and reiterate suffixation to the primitive base, in expressing a wealth of diminutive meanings (not even half of these are found in English), and in productively forming diminutives on a large variety of bases.

English, on the other hand, is more restricted in its (synthetic) diminutive expression. The language demonstrates less flexibility in the repertoire of word classes applied in the formation of diminutives. The only really productive diminutive class, both of primitives and diminutive derivatives, is the grammatical class of nouns. Adjectival and verbal diminution occur, but to a much smaller extent than nominal formation (and much less productively than Spanish). The use of recursive suffixation occurs in English, but is largely restricted to the domain of child language (e.g *dadd-y-kin-s*).

That said, however, much of the research referred to in this chapter to describe English diminutives is based on stereotypical assumptions about the status of the synthetic manifestation in the language (an exception is e.g. Schneider (2003)), which may have fabricated such an unfavorable view of the use diminution in English. The research presented in chapter 4 will try to uncover some facts about the "real" empirical state of English diminution, to get rid of the common myth of unproductiveness in the English language.

More recently, research on diminutives has tended to focus on the pragmatic uses of diminutives, e.g. Dressler and Merlini Barbaresi 1994 and Augustyn and Gniecka 2011, who claim that this approach is essential to the study of diminutives as they can only be studied in context. The present study looks at diminutives from a quantitative approach, aiming to study frequencies of formation and the distribution of semantic denotations, in order to attest the productive diminutive patterns existing in the two languages.

3 Method and material

3.1 Introduction

In setting out to describe and explain the occurrence of diminutives in language, several considerations need to be addressed in order to assure that the method employed and the material chosen will provide satisfactory answers to the hypothesis in question. The following chapter will discuss the use of contrastive analysis as a method in contrastive research (3.2), and similarly the use of corpora as material in a contrastive study (3.3).

This study acknowledges the distinction between corpus-based versus corpus-driven linguistics, as introduced by Togini Bonelli (2001), and argues that the present study is a corpus-based one. McEnery and Hardie (2012: 6) state that "corpus-based studies typically use corpus data in order to explore a theory or hypothesis, typically one established in the current literature, in order to validate it, refute it or refine it" (2012: 6). I argue that the description provided here by McEnery and Hardie (2012: 6) nicely describes my study's use of corpora. The present contrastive study aims to test the hypothesis of whether it is the case that synthetic diminutives are more frequent in Spanish than in English, through investigating authentic empirical language from two corpora.

Corpora open up ways in which we can investigate quantitative aspects of language quite effortlessly, where large amounts of empirical data can be studied, revealing e.g. frequencies of use, which I believe is essential for the topic of synthetic diminutives in English at this stage, as most research to date has been based on assumptions and findings gathered through introspection-based and dictionary-based research.

The choice of corpora is here of great importance, as it directly affects the validity and scope of the results. The two corpora, the *Corpus of Historical American English* (*COHA*) and *Corpus del Español* (*CDE*), have been selected here for similarity of composition (both created by Mark Davis), i.e. their large sizes (400 million and 100 million respectively), the wide variety of language types represented, the similarity of time periods covered, and similarity of corpus interfaces and search functions, all four factors contributing and ensuring the most important criterion of comparability in a corpus-based contrastive study. Noteworthy, the two most important criteria for the decision of corpora in my study were to ensure comparability and to provide a quantitative basis for my research, i.e. find two large-sized corpora that could be compared. Less important was the diachronic perspective of diminutive use which these two corpora also provide, as I have reason to believe that e.g. the -

-ie/-y/-ey-suffix has enjoyed recent popularity in Australian English, and is would be interesting to see if this had started to happen in American English as well, by looking at a corpus of synchronic American English instead[9].

3.2 **Method**

In the comparison of a linguistic feature in two languages, a feature which may very well be wholly different in the respective languages (e.g. English and Malay), it is crucial to ensure a systematic comparison of the phenomenon as this will produce the most viable results of the comparison. As such, the method of contrastive analysis will be applied here, selected for its appropriateness in corpus-based contrastive studies. This study aims to provide the methodology and framework principles of the method of contrastive analysis as proposed by Lefer (2011), which is based on James (1980), Krzeszowski (1990) and Chesterman's (1998) approaches to contrastive analysis as a methodology.

3.2.1 Contrastive analysis: A three-step approach

Lefer (2011: 650) proposes a three-phase analysis of the phenomenon in question. The first step consists of a description of the topic, i.e. a separate description of the topic in each language, which is essential in cross-linguistic studies in order to enable a comparison. Secondly, the descriptions of the topic are juxtaposed, which thus should establish a common ground for the comparison, a *tertium comparationis*, where a specific element of comparison is selected, which will be the starting point for the discussion. Thirdly, the actual analysis takes place, i.e. the analysis proper, which is a contrastive analysis of the element selected in stage two. A more detailed description of each of the three step of the analysis procedure will be presented consecutively. Stage two will also include the steps taken to acquire a *tertium comparationis* for the topic of the present study. In chapter four, the method will be applied throughout the analysis and discussion of diminutive frequencies and semantic patterns as revealed through the current corpus investigation.

[9] An initial thought, which had some impact on the choice of corpora as well, was to include a diachronic aspect of diminutive use in this study, as e.g. to trace the development of the *-ie/-y/-ey*-suffix the last two hundred years. Lack of time and space however made realise that the diachronic aspect here had to go.

Description

The aim of the description is to provide "data-induced" contrasts, instead of "description-induced" contrasts, i.e. authentic linguistic contrasts rather than reflecting the different uses of methods by the linguists, as argued by James (1980: 63 – 64). To accomplish such a comparison of "data-induced" contrasts, the same "model" of description must be applied in the account of both languages, i.e. the same variables in the respective languages must be described (James 1980: 63). Regardless of the theoretical framework applied, it is often a good idea to apply different terminology for the phenomenon in the two respective languages, i.e. the terminology that best describes the phenomenon in that particular language, as some aspects of language are very-language specific. Here, some confusion might arise, as descriptions which apply different theoretical frameworks have a tendency to become quite unclear, which can result in unfitted objects of comparison (Lefer 2011: 651).

The formula for success is then, as argued by James (1980: 64), to "describe L1 and L2 data independently, using the models which yield the fullest descriptions of either language, and then translate these two descriptions into a form that is model-neutral" (1980: 64). This is the idea that has been adopted in the current study's description of the topic. Thus, in chapter two, where the diminutive expression is accounted for, different terminology has deliberately been applied to discuss the topic of e.g. suffixes in the two languages. The Spanish suffixes are described as "emotive" and "non-emotive" (semantics, cf. e.g. Lang 1990) while the English ones, in contrast, have received the captions "neutral" and "non-neutral" (phonology, cf. e.g. Katamba 2005), thus respecting the language-specific features for each language.

In the description phase, yet another important factor is the use of basic notions such as e.g. derivation, derivational suffix, etc., that express the same concept in two languages, as no cross-linguistic definitions have yet been formulated. Lefer (2011: 652) states that the process of 'derivation' defined as 'an affix attached to a base' is relatively imprecise. We cannot immediately assume that the use of "derivative", "affix", "compound", etc. refers to the exact same linguistic process (James 1980: 167; Chesterman 1998: 32). If such assumptions are made, there is the very possible outcome of carrying out a comparison that is, in fact, incommensurable (James 1980: 167). I have for my own contrastive analysis in chapter 4, looked into what English and Spanish grammars refer to the process of derivation, and similarly, which semantic meanings diminutives in the two respective languages are said to have. These accounts have been placed either in the 'description'-stage of juxtaposition-stage, depending on where I found it most naturally.

Juxtaposition or establishing of a *tertium comparationis*

The next step of the contrastive analysis is to establish a common ground for the comparison, a *tertium comparationis* (TC), i.e. to find a factor, a shared characteristic that links the two elements in comparison, which then, will allow for a comparison of the cross-linguistic elements (*OED* 2016). Goddard and Wierzbicka (2008: 205-206) define the *tertium comparationis* as any "descriptive parameters which are stable and language-neutral", which, similarly, are independent of any language peculiarities (Goddard and Wierzbicka 2008: 206). Chesterman (1988: 29) and James (1980: 168–169), furthermore, elaborate that the TC acts as a "benchmark" for the comparison, emphasising, nonetheless, that this shared characteristic need not be identical in the two languages.

A problematic issue in the field of contrastive linguistics is the general trend to refrain from assigning a specific TC in the methodological framework of the studies. Indeed, Lefer (2011: 665) reports that few studies practice the method of contrastive analysis in full, referring to her own scrutiny of 70 contrastive studies from *Languages in Contrast* where only a mere 11% of the studies had a clearly developed TC. James (1980:169) acknowledges this problem, arguing its utmost importance as the differences discovered are only significant if studied against that same background, which is best achieved through the establishing of a common ground for the comparison (James 1980: 169). Similarly, Krzeszowski (1990: 15) backs up these claims, arguing that establishing a solid TC makes all the difference between a good and less good investigation, as it is the one element that the differences and similarities will be based on.

It should, however, be mentioned that a possible explanation for the absence of TCs in contrastive research to date, may be the lack of a precise definition of the concept. In fact, there is no existing formula for a TC that can be applied to all areas of contrastive linguistics, and thus, selecting the right one for the particular study is rather difficult (Lefer 2011: 655).

Previously applied TCs in affixation

In their study of contrastive prefixation, Lefer and Cartoni (2011: 97) apply a semantic type of TC, which proposes a range of common cognitive categories for the derivational process of prefixation in English, French and Italian, (namely "location, evaluation, negation, quantity, modality and inchoativity"). In derivational suffixation, however, cognitive categories would be incompatible as a TC. Suffixes are primarily applied in derivation to change the grammatical class of the base, and furthermore, the denotations are much more general, having more grammatical functions than prefixes (Lefer 2011:

665). Consequently, for the present study, i.e. for English and Spanish diminutive suffixation, another type of TC is required.

Lefer (2011: 665) proposes that to establish a TC for contrastive suffixation, the grammatical class of the primitive and the grammatical class of the derivative must first be identified. As the present study only looks at the diminutive derivative (not the primitive), this should then be the starting point for the establishment of a TC.

TCs for the present study

As the present study investigates and contrasts two variables of diminutive suffixation, namely frequencies and semantic diminutive patterns (see 4.2 and 2.3), two TCs must be established, one for each research question. Before the TCs can be developed, however, a few ground facts about the English and Spanish languages must be acknowledged in order to make sure that the element contrasted, namely diminutive suffixation, is the same process in both languages. In English, suffixation is typically described as "the creation of new lexical items by adding affixes" (Katamba 2005: 54). The production of new lexis can occur through the addition of a derivational suffix. Inflectional suffixes, on the other hand, add grammatical information, producing instead new word forms. Katamba (2005: 54 – 56) distinguishes, however, between "derivation by affixation" and a broader definition of derivation which includes all derivational word-forming processes, i.e. compounding, conversion, etc. The present study will apply the more narrow definition of "derivation by affixation".

The *Nueva gramática de la lengua española* (2009: 23) defines this word-formation process in Spanish as follows:

> en la derivación se une una base léxica y un afijo
> [in derivation a lexical base and an affix are put united]

It is evident, here, that both languages' definitions refer to the same word-forming process, i.e. one involving the attachment of an affix to a base. It can thus be stated that the process of derivation shares the same characteristics both in English and in Spanish.

Frequency TC

For research question one (RQ1) which looks into the frequencies of English and Spanish diminutives (see either 1 or 4.1), a suitable TC must be found, preferably a TC which can compare this variable in the most "neutral" manner. Since Spanish is in general more

productive in all aspects of diminution, the question is whether it is possible to find an area of English diminution which at least is relatively productive.

One area of diminutive production where English demonstrates a rather frequent pattern of diminutive production is in the grammatical class of nouns (e.g. Quirk *et al*. 1972). Diminutives that commonly occur in such formations are *kitchenette* (from *-ette*) *booklet*, *streamlet* (from *-let*) *duckling* (from *-ling*), and *daddy*, *auntie* (from *-ie/-y/-ey*) (Bauer 1983: 221; *COHA* 2016; Quirk *et al*. 1972: 994-995).

Similarly, and more profoundly, Spanish is also productive in the formation of nominal diminutives, as stated by *Nueva gramática de la lengua española* (2009: 626): Similarly, and more profoundly, Spanish is also productive in the formation of nominal diminutives, as stated by *Nueva gramática de la lengua española* (2009: 626):

> Se llaman APRECIATIVOS los sufijos que se añaden a numerosos sustantivos y adjetivos, y ocasionalmente también a otras clases de palabras […]

Thus, the TC that will be applied in research question one, which tests the contrasting frequency of the diminutive category in English versus Spanish, is the grammatical category of nouns.

Semantic TC

Research question two (RQ2) (see 1 or 4.1) will investigate the semantic patterns of diminutives in English and Spanish, which thus requires the establishment of a semantic foundation or TC for the contrastive analysis

Languages tend to apply different taxonomies in the accounts of suffixes, which reflect the properties that best describe/most appropriately describe the way the suffixes behave in that language (Katamba 2005: 58). English suffixes, for instance, are often grouped according to the phonological changes they invoke on the base lexeme, i.e. as English suffixes typically change the stress of the base. Spanish suffixation, on the other hand, typically groups suffixes according to the type of denotation expressed, thus commonly dividing the suffixes into classes two different classes, i.e. emotive and non-emotive suffixes (Lang 1990: 91). Diminutive suffixes belong to this emotive group, as they typically express an emotional element (except for specialised diminutives). English suffixes could, for that matter, also be categorised according to meaning, but this grouping would in many ways be unsatisfactorily, as English suffixes primarily have grammatical functions (e.g. change the word class) and in

addition because the extent to which diminutive and e.g. augmentative suffixes are used is relatively small in comparison.

Regardless of the taxonomies applied in different languages, however, the core meaning expressed by diminutives is very much the same within the various languages. This point is demonstrated by Strang (1968: 138), who ascribes the following semantic properties to English diminutives:

> Diminutives are usually forms that have begun by meaning 'a small one of its kind' but have undergone a development whereby they come to express not merely an assessment of size, but also, or even exclusively, the speaker's response to small things, a response ranging from affection through condescension to contempt; we might even say that a diminutive is mature when it carries only this 'response'-meaning.

Spanish emotives, or appreciatives, which comprise diminutives, pejoratives and augmentatives, are characterised as follows:

> Se llaman APRECIATIVOS los sufijos que se añaden a numerosos sustantivos y adjetivos, y ocasionalmente también a otras clases de palabras, para expresar tamaño, atenuación, encarecimiento, cercanía, ponderación, cortesía, ironía, menosprecio y otras nociones – no siempre deslindables con facilidad – que caracterizan la valoración afectiva que se hace de las personas, los animales o las cosas […][10]
>
> (*Nueva gramática de la lengua española* 2009: 626)

These two semantic descriptions of the respective uses of diminutives in English and Spanish have some shared diminutive senses:

smallness ~ tamaño 'size', atenuación 'reduction, lessening';

affection ~ cercanía 'closeness, nearness';

condescension/contempt ~ menosprecio 'scorn, contempt'[11]

These three meanings will be applied as a TC for the diminutive comparison of semantic meaning in research question two (RQ2). In addition, both languages' diminutives express ambiguity, and thus, this will constitute a fourth semantic category. A fifth category will be

[10] Translation: [They are called APRECIATIVES the suffixes which are added to numerous nouns and adjectives, and occasionally also to other word classes, to express size, reduction, increase/emphasis, closeness, praise, politeness, irony, contempt and other notions - not always easily definable – which characterize the emotional evaluation/assessment directed at people, animals or things […]]

[11] Definitions for the Spanish lexemes are taken from *WordReference* (2016).

included as well, because of its extremely frequent use in Spanish, namely specialised diminutives. Specialised diminutives are non-existent in the English diminutive descriptions I have come across. The category will, however, be included in the contrastive semantic analysis. There might be no findings of this diminutive type, but this particular meaning might not have been looked for in previous empirical research of English diminutives.

Contrastive analysis proper

The third and final step of the contrastive analysis consists of the contrastive analysis proper. Here, the elements can finally be contrasted, as descriptions and TCs now exist, producing a proper theoretical background and foundation for the contrastive comparison.

The analysis proper can produce three different outcomes of the analysis (Lefer 2011: 656). Possibility A is that the contrasted element appears to be largely similar in language X and Y. The second possibility (B) is that the element in language X diverges in some degree from that of language Y. Lastly, there is the possibility (C) that the element in language X is non-existent in language Y (Krzeszowski 1990: 37-38).

To discover which of the possible outcomes that is true for my data, a few statistical testing methods will be applied for the semantic diminutive analysis in RQ2. RQ1 will mostly be tested through calculations of normalised frequencies (per millions words) in the two languages.

3.2.2 Previous contrastive research

The types of studies reviewed here are those belonging to the field of contrastive word-formation, which consists of three separate linguistics fields, i.e. morphology, lexicography and contrastive linguistics. To date, this particular field is an under-researched field of study (Lefer 2011: 647). This section is largely a summary of Lefer (2011) which provides an overview of the main trends in the field.

Lefer (2011) examines a large selection of contrastive studies from the last fifty years (1960–2010). The selection consists of a mixture bibliographies, conference programmes, journal articles and handbooks. The complete bibliographies of contrastive linguistics (e.g. Hammer and Rice 1965, Sajavaara and Lehtonen 1975) disclose that only 1% of the 2,200 studies included concern contrastive word-formation, which according to Lefer (2011) well represent the "current trends of the disciplines under scrutiny" (2011: 647-648). Another set of bibliographies, i.e. Stein (1973), Beard and Szymanek (1988) and Scalise (2005), demonstrate

that contrastive research still mainly appears to consist of typological and monolingual studies (Lefer 2011: 648). These typology studies have mainly addressed the topics of grammar, syntax and phonetics.

Importantly, though typology studies provide important knowledge about languages, they are not contrastive studies, but comparative studies, i.e. descriptive accounts of languages. Typology (or comparative) studies have traditionally served to establish historical connections and similarities between languages (cf. e.g. Terrell and Salgués 1979; Whitley 1986; Lipski 1990: 250). Few of these treat morphological questions, though there are some (e.g. Savickiene & Dressler 2007), which indeed are valuable for the establishing of the *tertium comparationis* for contrastive word-formation studies, as they describe typical morphological features for different languages (Lefer 2011: 648).

Up to the present, contrastive word-formation have mainly dealt with the topics of suffixation, compounding and prefixation, however restricted the quantity of the research is. Lefer (2011: 649) states that of the three word-forming processes, suffixation is by far the most researched represented by works such as e.g. Bertrand 1986, Booij & Lieber 2004, Ascoop & Leuschner 2006, Heyvaert 1998 and Kin Quah 1999. Some studies addressing the topic of verbal compounding are e.g. Bauer 1978, Paillard 2000, Arnaud & Renner 2008, Gast & Hüning 2008 and Guevera & Scalise 2004, and in addition, a few focus on prefixation, e.g. Andor 2005, Fedorowicz-Bacz 1977, Funk 1986 and Lefer & Cartoni 2011.

Unfortunately, much of the existing research on contrastive word-formation is still largely rooted in the introspection-based and dictionary-based traditions. Here, invented examples and words occurring in the dictionaries are employed as evidence for the occurrence of e.g. a particular type of suffix. In fact, only one sixth of the word-formation studies Lefer (2011: 650) examined apply corpus as the main source of material (cf. e.g. Cvilikaitė 2007).

The recurring languages in the contrastive studies in Lefer's (2011) overview were, as a rule, the three languages English, German and French, which appeared in various combinations with a variety of other European languages (e.g. Hungarian, Polish, Serbo-Croatian, French, Dutch, Danish, Swedish, Spanish, Lithuanian, Bulgarian, Russian, Portuguese, Finnish, Romanian, Italian, Modern Greek), in addition to some studies where a non-European language was part of the study (e.g. Arabic, (Modern) Chinese, Malay, Japanese).

Only two studies here concern the variables pertaining to the present study, i.e. English-Spanish word-formation (Rabadan, Labrador & Ramon 2005) and diminutives (Sifianou 1992).

English-Spanish contrastive studies

The few studies that I have come across which contrast an aspect of word-formation in English and Spanish, (which are not typological studies) are Williams (2006), in addition to the two just mentioned in 3.2.3 (i.e. Rabadan, Labrador and Ramon (2005) and Ramon (2006)). None of these contrastive studies, however, concern diminution. One addresses adverbs in *-ly*, the second nominal modification, and the third a quantitative model for translation studies.

3.3 Material

3.3.1 Introduction

This section will address a few factors that are important in the application of corpus as material in research. In addition, the two corpora that have been applied in the contrastive analysis here will be introduced (3.3.3 and 3.3.4). The application of corpus as material in contrastive linguistic research has opened many new doors to the possibilities of what is achievable in research. As Stubbs (2002) acknowledges:

> Corpus methods can organize huge masses of data, and make visible patterns which were only, if at all, dimly suspected. In giving access to new data, the technology opens up research topics which were previously inconceivable. We now have facts about language use which no amount of introspection or manual analysis could discover […].
>
> (Stubbs 2002: 221)

Thus, through the advent of electronic corpora, contrastive linguistics, in particular, has experienced a face-lift as language has become so much more easily accessible for research and scrutiny. Large amounts of empirical data, i.e. actual produced language existing in quantities (millions of words), can be studied, revealing characteristics, patterns and trends about language, which dictionaries and introspection can only dream of.

3.3.2 Comparable corpora

To ensure the quality of a corpus-based study, it is important to ensure that the two (or more) corpora are comparable. It is especially important in the comparison of corpora that have not been *designed* for comparison, as is the case with the two corpora of the present study, which reflect texts of two different languages. Rayson and Garside (2000) states

that there are two types of corpus comparisons, i.e. "comparison of a sample corpus to a larger corpus" and "comparison of two (roughly-) equal sized corpora" (2000: 1). The present study applies the second type of comparison, namely a comparison of two large corpora, as the study aims to discover whether some features are different in the one corpus compared to another. As such, particularly the issues of homogeneity within the corpora and comparability of the corpora need to be considered (Rayson and Garside 2000: 1). This section will briefly discuss these two factors.

In order to ensure homogeneity within the corpora of comparison, the composition needs to be controlled for uniformity, i.e. that roughly the same composition of genres is reflected in both corpora. Otherwise, the results of the corpus-analysis may reflect only the differences of the corpora's contents, i.e. that one of the two contained more informal language and thus demonstrated larger findings of this type (Kilgarriff 1996).

The issue of comparability of corpora concerns the assurance of likeness between the contents of two corpora, i.e. that like is being compared with like. In this respect the sampling method used to compile the corpora is of importance, which means that e.g. the same "randomised methods of sample selection" (Rayson and Garside 2000: 2) are employed in the compiling. The corpora selected for the present study are both the work of Mark Davis, i.e. *COHA* and *CDE*, and as such, I hope that that should logically entail the comparability.

3.3.3 *Corpus of Historical American English* (*COHA*)

The *Corpus of Historical American English* (*COHA*) was released in 2010 and is one of the latest additions to Mark Davis' repertoire of linguistic corpora. The corpus consists of more than 400 million words distributed on more than 100,000 individual texts. It is a historical corpus covering a larger period of the English language (1810–2009).

There are four genres in the corpus, each representing a different type of language, allowing the linguist to investigate language and language changes resembling that of the real world. Those four genres are Fiction, Magazine, Newspaper and Non-fiction. All these genres are balanced across the decades (1810-2009) which means each genre in quantity of words is equally represented for each decade. Fiction accounts for 48-55% of the total in each decade, while the three other genres are represented equally, though in lower percentages, across each decade.

The only category missing in *COHA* that is represented in the Spanish parallel corpus, *CDE*, is spoken language. This is a little unfortunate, first, in terms of ensuring the comparability of the two corpora, and secondly, because diminutives tend to occur in the more informal parts of language. One fourth of the data from the 1900s in *Corpus del Español* (see 3.3.4) is spoken language (which is one eight of the total data that this study applies from *CDE*, i.e. 1800s and 1900s together, as the data from the 1800s consists only of fiction). That said, the Fiction genre in *COHA* accounts for roughly 50% of the total in each decade (1810s-2000s), which means that half of the data in *COHA* is fiction "texts". The Fiction genre is made up of books (e.g. novels, prose, drama, e-books, scanned books, social history articles), movies and play scripts. The language found in these sub-genres represent spoken language quite well, as i.e. indirect speech, dialogue in movies, drama dialogue, etc., which thus equals quite well the Spanish oral data (interviews, etc.).

3.3.4 *Corpus del Español* (*CDE*)

Corpus del Español was first published in 2002, but in 2007 it underwent a large revision where, among other things, its size was more than doubled. It now contains 100 million words consisting of texts covering an extensive variety of genres representing real language, i.e. the genres of Literature (Fiction), News, Academic and Oral language. The Spanish represented are both peninsular and of Southern-American. It, too, is a historical corpus dating form 1200s to 1900s.

The present study will only apply the data from the periods of 1800s and 1900s from this corpus, in order to mirror the diachronic scope of *COHA* (1810–2009), the other corpus used in this study. The data from the 1800s in *CDE* are exclusively novels. The 1900s material is more varied, the texts evenly distributed among genres of Fiction, Newspaper, Academic and Spoken language (Davis 2010: 140). The Fiction genre contains novels and short stories, the Newspaper genre consists mainly of news articles, the Academic texts is made up of a range of scholarly articles and, finally, the Oral genre which is chiefly a mix of interviews and transcriptions, more than five million words deriving from transcripts of spoken conversation, often from a political setting (Davis 2010: 140).

The revision of *CDE* which was completed in 2007 has made the corpus data more available for comparison with other languages (and, of course, monolingual studies), as the tagging of the lexemes was improved, and various registers built in to the search interface,

both factors adding to the enhanced applicability of the corpus in contrastive studies, since large corpora of a selection of other languages already have these advanced functions.

3.3.5 Retrieving language data from *COHA* and *CDE*

In this section, the process of retrieving and selecting relevant data from the two corpora for the contrastive diminutive discussion presented in this study will be accounted for. First, options for different search functions in the interfaces of *COHA* and *CDE* will be briefly stated. Secondly, the process of designing search strings for the retrieving of the appropriate language from each language will be discussed. Due to a language-specific difference between English and Spanish, the double amount of search strings were required to retrieve the same data in the Spanish corpus compared to the English corpus. The specification of features in the search strings were, however, the same.

Search functions

The interfaces of *COHA* and *CDE* are the same, both allowing for four different search modes, i.e. 'LIST', 'CHART', 'KWIC' and 'COMPARE'. Different variables and information are revealed about the lexemes or parts of speech (POS) by choosing a different function.[12] As I was interested in an overview or a list of lexemes and their corresponding frequencies, the 'LIST'-function was selected and is the function mainly used in the retrieving of data.

Finding the correct search strings

In order to find the correct search string for my purposes, i.e. one that would provide me with nominal diminutive lexemes, the suffix [SUFFIX] was entered into the slot, in addition to a tag [ALL NOUNS], and lastly also a full stop between the two bracketed specifications/items, indicating that the suffix and noun is the same item, not two adjacent or lexemes. These two specifications would provide me with a list of nouns that ending in the suffix specified in each case, but also words ending in the same letters as a suffix.

For the English data, it was often sufficient to put square brackets around the suffix to indicate that I wanted to see [ALL FORMS] of the lexeme ending in a suffix, both singular and plural forms, e.g. that [*ie] would give me words like *lie* and *lies* alike. However, to be

[12] See http://corpus.byu.edu/coha/

certain that all plural forms occurred in the search results, each diminutive that had been selected, was in the end searched for separately in three different ways, applying the singular form of the diminutive in one search, the plural form in the second search, and lastly the diminutive in square brackets, to discover whether any of the results differed (e.g. 1: 'daddy'; 2: 'daddies'; 3: '[daddy]'). More than once, the raw frequency varied when applying one of the alternative search strings. The bracketed word should provide you with all forms of the word, but apparently, as previously also detected in corpus-studies undertaken, the tagging is not fool proof. Here, to ensure that the frequencies did not differ depending on which one or two search string I applied for each separate diminutive, I chose to apply 1 and 2 and leave the bracketed one. Normally, one estimates the fault rate to be roughly 97%.

For the Spanish data, another problem occurred, in addition to the issue of including plural forms, i.e. that the square brackets failed to provide both feminine and masculine forms of the lexeme, in addition to plurality. Thus, a double amount of primary search strings has been used to retrieve the data from the Spanish corpus. Apparently, the compiler or tagger of CDE has treated feminine and masculine suffixes and lexemes separately (e.g. [señorito] will not provide you with the word forms *señorito*, *señoritos*, *señorita* and *señoritas*).

The full list of primary search strings applied in the retrieving of English diminutives can be found in table 3.1 and corresponding strings for the Spanish diminutives in table 3.2:

Table 3.1: Primary search strings applied in the retrieving of English data

English data	
Suffix	**Search string(s) in *COHA***
-ie(s)/-y/-ey(s)	[*ie].[nn*], [*y].[nn*]
-let(s)	[*let].[nn*]
-ling(s)	[*ling].[nn*]
-ette(s	[*ette].[nn*]

Table 3.2: Primary search strings applied in the retrieving of Spanish data

Spanish data	
Suffix	**Search string(s) in *CDE***
-ito/a(s)	[*ito].[nn*], [*ita].[nn*]
-ico/a(s)	[*ico].[nn*], [*ica].[nn*]
-illo/a(s)	[*illo].[nn*], [*illa].[nn*]
-ín(es)/ina(s)	[*ín].[nn*], [*ina].[nn*]
-ete/a(s)	[*ete].[nn*], [*eta].[nn*]

Altogether, the primary search strings for both languages count fifteen. These fifteen search strings each provided me with the lexemes ending in any of the fifteen variants of suffixes, i.e. lexemes ending in *-ie(s)/-y(s)/-ey(s), -let(s), -ling(s), -ette(s), -ito/a(s), -ico/a(s), -illo/a(s), -ín(es)/ina(s)* and *-ete/a(s).*

3.3.6 Process of sorting relevant data from the irrelevant

After having retrieved the list of lexemes procured from the corpus searches, the task was then to sort and rule out diminutives from non-diminutive lexemes. Some words proved more difficult than others to categorise. Throughout the sorting process, the *Oxford English Dictionary* (*OED* 2016) and *Diccionario de la lengua española* (*DRAE* 2016) have been consulted to check the meanings of the lexemes, so as to make sure that no diminutive lexeme was left out of the discussion that should have been included. Especially in the determination of the Spanish data, as Spanish is my third language, *DRAE* (2016) was consulted rather frequently.The lexemes that explicitly expressed a diminutive meaning, and were formed on an authentic diminutive suffix, were swiftly ruled diminutive. Instances of lexemes such as *lie* and *prairie*, which simply terminate in the same letter combinations as a diminutive suffix, have of course been ruled non-diminutive.

This is quite time-consuming work, and in time, I hope that Mark Davis' corpora, or corpora in general (I assume some have this function), will acquire a tagging function which allows searching for various meanings, e.g. diminutive meaning.

4 Discussion and results

4.1 Introduction

This chapter aims to explore the state of contrastive English-Spanish diminutive formation by addressing two fundamental concepts of diminution, i.e. lexeme productivity and the distribution of semantic denotations. The discussion of this chapter will build on previous research and basic assumptions about diminutives in English and Spanish (see chapter 2). As previously discussed, the two languages have very different traditions of applying synthetic diminution in language. Spanish, a classical synthetic language flourishes in creative and productive diminutive lexis, whereas English, on the other hand, prefers analytic diminutives, applying synthetic diminutives more rarely than Spanish. As such, it will be interesting to see how diverging the diminutive manifestations in English and Spanish are through the two corpus-analyses that will be carried out in this chapter. The observations made in the two analyses will, hopefully, help to establish some ground facts about the two different manifestations of diminution in English and Spanish based on empirical evidence.

4.1.1 Structure of chapter

The logic applied in this chapter to test the two diminutive characteristics can be gathered from figure 4.1, which states the hypothesis of this diminutive investigation, which is the root of all efforts made in this study, and the research questions used in order to either confirm or reject the hypothesis:

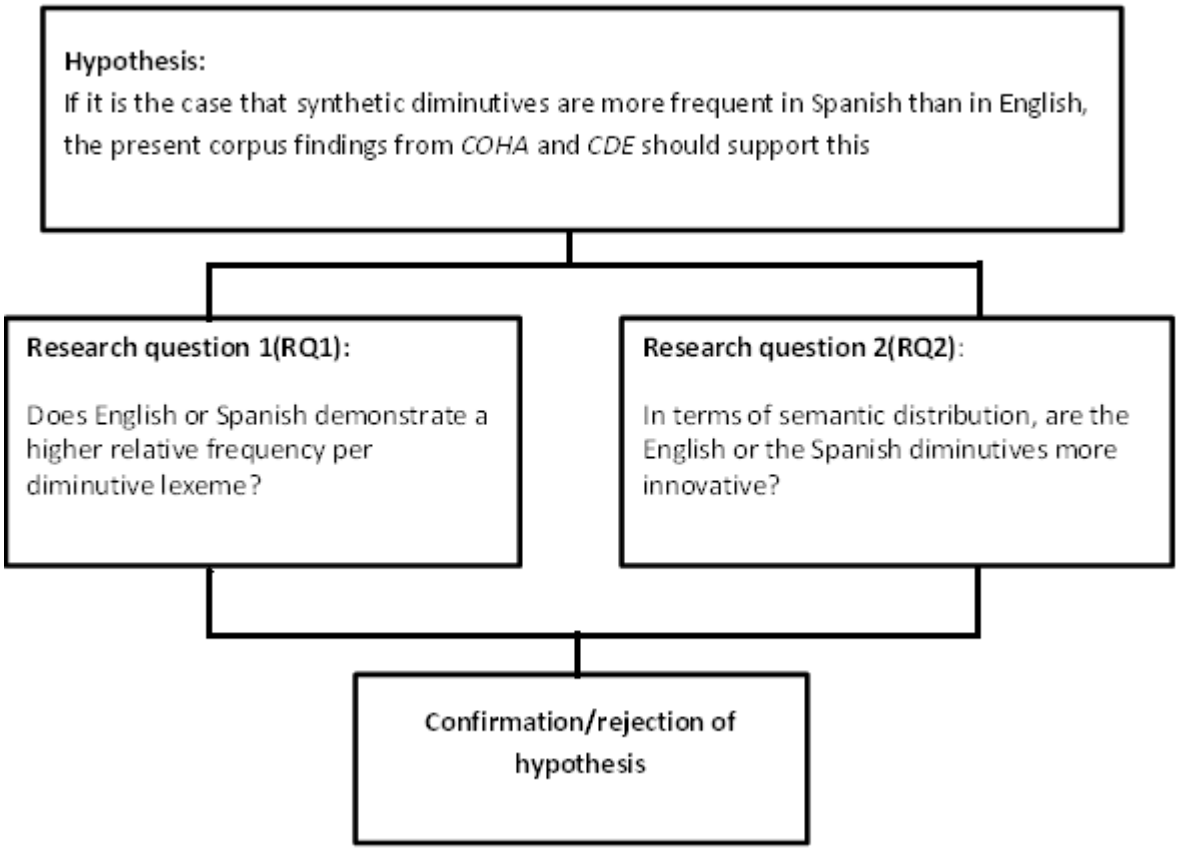

Figure 4.1: Build-up of this chapter

RQ1 will be addressed through a diminutive frequency analysis which will take place in 4.3. Similarly, RQ2 will be addressed in a semantic-diminutive distributional analysis of 4.4.

However, before I look into these two fundamental concepts of diminution, a few preliminary questions need to be addressed (4.2). This section of 4.2 will look into some questions concerning the analyses to come, i.e. the decisions made regarding the choice of diminutives to be analysed. Firstly, the suffixes selected for analysis will be discussed. Secondly, language types and language varieties chosen for study will be stated. Thirdly, the process of identifying diminutive meaning and sorting diminutive lexemes from non-diminutives will be accounted for, and a few difficult cases will be illustrated.

The second and third sections (4.3 and 4.4) will concern the two diminutive analyses of the present study, which naturally constitute the main parts of this chapter. These two analyses will test the productivity and the semantic distribution of five diminutive denotations in English and Spanish diminutives. The frequency analysis looks at contrastive diminutive frequencies of English and Spanish, testing the productivity of English and Spanish synthetic

diminutives. The semantic distribution analysis will look at the productivity of the various semantic distribution patterns of diminutives in English and Spanish. The two discussions will mainly be based on the comparison of normalised frequencies of the lexemes. The semantic analysis will, in addition, also include some statistical testing, i.e. primarily in the form of the statistical method of log likelihood, which will be discussed in 4.4.1. The two sections will both end with a short summary of the respective findings of the analyses and point to diminutive trends observed in the two languages. Together the two diminutive analyses should be able to confirm or reject the hypothesis' claim, which tests whether Spanish (really) is more productive than English in synthetic diminutive formation.

The answers to the two research questions will, I hope, contribute to a better understanding of some fundamental characteristics about diminutives, namely diminutive lexeme productivity and semantic diversity expressed by the diminutive category, as tested in a typical emotive language, Spanish, and a less synthetically expressive language, English.

In addition, interesting findings about the different uses of diminutives in English and Spanish could potentially/possibly be applicable to other types of research, e.g. translation studies, learner language or second language acquisition. At the very least, it offers an informative read of a topic that has rarely been addressed from this angle in research to date.

4.1.2 Build-up of the analyses

The two diminutive analyses will each be addressed through the method of contrastive analysis as proposed by Lefer (2011: 650)[13] (see 3.2.1). The application of the method of contrastive analysis is of utmost importance here, as it establishes a framework for the comparison, ensuring, among other things, that like is compared with like (see 3.2.1). This is particularly important in the discussion of topics such as the present one, which currently lack a framework for the discussion (cf. James 1980; Krezszowski 1990; Chesterman 1998). Consequently, the two diminutive analyses will follow three steps of comparison, i.e. description, juxtaposition and an analysis proper (see 3.2.1). The description phase will provide two accounts of the diminutive variable in question, one concerning English and one for Spanish. The second stage will carry out a juxtaposition of the two descriptive accounts, which thus provides a *tertium comparationis* for the comparison. Finally, the third stage will

[13] Based on James (1980), Krezeszowski (1990) and Chesterman's (1998) contrastive methodological frameworks.

concern the actual contrastive analysis proper of the diminutive variable, which makes out the main part of the analysis. Here, corpus findings will be presented, analysed and discussed.

4.2 Preliminary questions

Before setting out to analyse the corpus data from *CDE* and *COHA*, and present the findings from the two respective diminutive analyses, a few factors need to be considered regarding how the corpus data came to be. First, a short discussion of the selection of diminutive suffixes included in the analyses will take place. Secondly, the language types and language varieties chosen for study will be discussed. Thirdly, the process of identifying diminutive meaning and sorting diminutive lexemes from non-diminutives, with emphasis on the difficulties that arose/presented themselves, will be accounted for.

4.2.1 Selection of diminutive suffixes

The investigations into diminutives carried out in this chapter have focused on a range of diminutive suffixes which, as indicated by previous research, are considered to be a representative selection of diminutive suffixes in English and Spanish. The suffixes have been selected carefully, as the assigned repertoire of suffixes will, to a large extent, be defining for what the findings will show. Five suffixes have been selected for the collection of the Spanish data:

-ito/a(s), -ico/a(s), -illo/a(s), -ín(es)/ina(s), -ete/a(s)

and four suffixes for the English data:

-ie/-y/-ey(s), -ette(s), -let(s), -ling(s)

Together these nine diminutive suffixes will demonstrate the trends of diminutive formation, in terms of productivity and meaning types, during the last two hundred years (c.1800–2000). A list of the primary search strings adopted for each of the suffixes has been given in 3.3.5.

4.2.2 Language types and language varieties looked at

Interestingly, an area of language which has received little attention in English diminutive research is informal language. This is quite odd, because this is one of the areas of language

where diminutives are most prone to appear. Wierzbicka (1985), for instance, reports extensive use of diminutives and so-called embellished clippings[14] in Australian English, a variety of English which we know to have a more relaxed style of speaking than e.g. British English. Diminutive research has often suffered from the difficulty of obtaining authentic speech, especially in quantities (see e.g. Schneider 2003: 69). In this respect, corpora provide an excellent opportunity to research this particular area of language, as amounts of naturally occurring language have been recorded or collected, and thus is represented mainly by the genres of Spoken language and Fiction (novels frequently apply indirect speech and movie scripts), which are well represented in the two corpora applied in the present study. More explicitly, the investigation of informal language can, if the results are positive, challenge the many generalisations which depict diminutive formation as poor in English.

This study will investigate diminutive formation in American English (AmE) and all dialects of Spanish (SpALL), both peninsular and Southern American varieties. The choice of these two varieties is to some degree coincidental, as the two most suitable corpora for comparison happened to be *COHA* and *CDE*.

4.2.3 Diminutive type under investigation

The type of diminutives that will be tested in the two diminutive analyses is nominal diminutive derivatives (base + diminutive suffix = nominal diminutive). This is the *tertium comparationis* (TS) for the contrastive discussion. The choice of TC was motivated by the fact that the nominal class is the most productive word class for diminutive formation, both for the type of primitive bases and diminutive derivatives, in both English and Spanish (see 3.2.1). Thus, the study of nominal diminutives will provide the most fruitful, feasible and fun diminutive discussion[15].

4.2.4 Assessment and selection

The nine diminutive suffixes used in the search strings for the analysis produced a long list of lexemes, both diminutive and non-diminutive ones, i.e. words ending in a pseudo-diminutive suffix, or simply just in the same letters as a diminutive suffix (e.g. *lie, prairie, música*

[14] Which some scholars treat as diminutives (e.g. Schneider 2003, cf. discussion in 2.7.1).

[15] As discussed in 3.2.1, though Spanish shows productive patterns in the diminutive formation of adjectives and verbs as well, English is only vaguely produce adjectival diminutives and even fewer verbal diminutives.

'music', *política* 'politics'). I selected 100-120 lexemes from each corpus/language (the relative frequency had at this stage become quite low in both corpora), which were a good mix of lexemes formed on various diminutive suffixes. These lexemes were then analysed further, to control and assure, in particular, (1) the derivational formation process of the derivative, (2) the denotative meaning expressed by the derivative, and (3) the etymology of the primitive/diminutive. The assessment demonstrated that a few of the selected lexemes did not have all these things in order, and were thus ruled non-diminutive. In the end I ended with a selection of 99 English and 84 Spanish diminutive lexemes. A particularly frequent problem in the assessment of both languages' lexemes was incomplete or uncertain etymology provided by the dictionaries. The consulted dictionaries have mainly been the *OED* (2016), *DRAE* (2016), *Collins Spanish Dictionary* (2005), and also alternatively *WordReference* (2016) for the Spanish lexemes.

In general, the assessment of Spanish diminutives turned out to be a little difficult as diminutive forms normally are not accepted into the dictionary because there are so many of them. Thus, other resources needed to be consulted. A reference work that has been a tremendous help in this respect is Gooch (1967), which lists a large number of diminutives attested in Spanish, together with their meanings in English and the diminutive type they belong to (e.g. diminutive-pejorative, specialised diminutive).

The following list of words (tables 4.1 and 4.2) are lexemes that have been ruled non-diminutive, and will not constitute a part of the analyses that will follow:

Table 4.1: Spanish non-diminutive lexemes[16]

abanico(s) ''fan; range; variety; derrick' borrico(s) 'she-donkey; fool; sawhorse, sawbuck' coqueta 'dressing table' gaceta 'gazette, newspaper; gossip' golondrina(s) 'swallow' ladrillo(s) 'brick; tile; block; a bore' orilla(s) 'gentle wind'	piquete(s) 'squad, party; prick, jab; small hole (in clothing)' rodilla(s) '(anat.) knee' sainete(s) 'one-act farce, one-act comedy' semilla(s) 'seed; source; baby, small child' tarjeta(s) 'small shield; card'

[16]Definitions of these lexemes have been founded in *Collins Spanish Dictionary* (2005), *DRAE* (2016), *WordReference* (2016) and Gooch (1967).

Table 4.2: English non-diminutive lexemes

baby(ies)	holly
Bessie|bessie 'hypocoristic of Elisabeth; best friend'	honey(s)|honie(s)[17]
dominie(s)	hurley(s)
donkey(s)	monkey(s)|monkies
gazette(s)	siebling(s)
genie(s)	tory(ies)

A few illustrative examples

This section will take a closer look at some common problems that occurred in the assessment process. Typically, lexemes that have been ruled non-diminutive involved cases of (a) denotative ambiguity, (b) suffix ambiguity and (c) derivational etymology. The following section will discuss a few difficult cases pertaining to these categories (a-c).

(a) Denotative ambiguity

In general, due to semantic change, lexemes typically express more than one meaning. Therefore, we can rarely rule a lexeme as diminutive because we know it to have a diminutive meaning (e.g. the primarily hypocoristic form *sweetie* has the additional denotations of 'e.g. sweet cake' and the attributive use of 'sweetie shop'). Consequently, lexemes that demonstrated such ambiguity had to be looked at more closely to find which denotation was used in the corpus. Thus, as the primary focus of this study is quantitative, only a random sample of 100-200 sentences was studied in each case, to discover the most frequent application of the lexeme in use. If few sentences expressed the diminutive sense, the lexeme was ruled non-diminutive (e.g. three instances of a lexeme expressing diminution is not enough to be included in an analysis of the most productive diminutives in English and Spanish).

[17] *Honey* is, like *darling*, often applied as a term of endearment, and thus belongs to the category of hypocoristics. However, it has several other meanings pertaining to the honey produced by bees. It is for this reason that it has not been included in the present study, as the scope of this study does not allow the researcher to go through the more than 11,000 hits for *honey(s)|honie(s).*

Spanish lexemes, in particular, tended to express several meanings of which only one was diminutive. Typically, the analysis of the random sample of the given lexeme demonstrated that the lexeme's most productive use was non-diminutive. One illustrative case is Spanish *semilla*(*s*), which has three different uses, i.e. (1) 'seed'; (2) 'source (of origin)'; and (3) 'baby, small child', e.g. *la semilla* 'the kids' (collectively) (informal) (*Collins Spanish Dictionary* 2005). The analysis of a random sample showed that *semilla(s)* was only rarely being applied in the diminutive sense of (3) here[18]. The majority of the cases referred to sense (1) of 'seed' as used in e.g. the domain of gardening.

(b) Suffixal ambiguity

A case of suffix ambiguity presented itself with the English suffix *-ling*. This English suffix has two allomorphic variants, i.e. in the form of Old Norse (ON) *-ling* and Old English (OE) *-ling*, each expressing a different sense. Only the ON variant forms diminutives. Therefore, all derivatives of *-ling* should not be assumed to be diminutives (e.g. *sibling, underling* are not diminutives).

The corpus data demonstrated nine cases of formations of the ON suffix *-ling*. I was able to assess them as diminutives based on the etymology descriptions of lexemes in the *OED* (2016). The majority of these diminutive formations express smallness or pejoration. Three diminutives denote the smallness sense of young of animals or plants, often occurring with diminutive suffixes, i.e. *yearling* 'an animal that is one year old', *sapling* 'a young tree' and *seedling* 'a young plant developed from a seed'. Another bulk are of the diminutive-pejorative type, most typically formed on "human" primitives, such as e.g. *foundling* 'a deserted infant whose parents are unknown' and *sapling* 'a young or inexperienced person'. Other diminutive-pejoratives from the corpus are *stripling, starling, dumpling, kindling, changeling* and *weakling*.

In contrast, the corpus data contained a number of frequent OE *-ling* formations as well, more precisely *sibling, darling* and *underling*. These three express the non-diminutive sense of 'a person or thing belonging to or concerned with (what is denoted by the primary n.)' (*OED* 2016). However, of the three OE formations, only *sibling* has been ruled non-diminutive. The lexemes *darling* and *underling*, on the other hand, have been assessed as

[18] One random sample of *semilla* (singular) and one for *semillas* (plural) were studied, which together comprised of 200 sentences.

diminutives despite their origins, because they express like typical diminutives. Diminutive *darling* 'one dearly loved' (from OE *déor* 'dear' + *-ling*) is commonly applied as a term of endearing address (*OED* 2016). As such, it is a hypocoristic diminutive, similar in use to *sweetie*, *cutie*.

The lexeme *underling* is a production of the adverb *under* 'into a position or state of subjection or submission' and OE *-ling*, but expresses the typical pejorative sense often occurring with the derivatives of the ON *-ling* variant (e.g. *weakling*). The lexeme *underling* occurs in various senses today, as in the nominal senses of (a) 'one who is subject or subordinate to another', (b) 'a branch, plant, etc., growing under, or less strongly than, another; a small or weakly plant, animal, or child' or in the adjectival sense of (c) 'undersized, small, weak' or (d) 'trivial, unimportant' (*OED* 2016).

The two cases, *darling* and *underling*, are special as the denotation of the primitive already expresses a "diminutive" denotation, here affection by OE *déor* 'dear' and pejoration by *under* 'into a position or state of subjection or submission' (*OED* 2016). It is only for this reason that they can be analysed as diminutives[19].

(c) Derivational etymology

Some lexemes appear very diminutive-like, but at closer inspection they are revealed to have been formed erroneously, in terms of the derivational process producing the lexeme. The majority of the lexemes belonging to this group occurred in the form of so-called embellished clippings (see 2.2.6). Embellished clippings add a diminutive suffix (or other) to an already truncated primitive form, as in e.g. *bestie* for *best friend* (*best-* + *-ie* > *bestie*). This group of lexemes differs from diminutives in form, in that the suffix attaches to a truncated word form, and in addition in meaning, as it is not necessarily smallness that is expressed, but rather solidarity towards the addressee (cf. Wierzbicka 1985; Dressler & Merlini Barbaresi 1994; Sussex 2004). As previously discussed, this group constitutes a rather controversial type of "diminutives", which this study claims is not diminutives (2.2.6). Some lexemes from *COHA* that are of this type are: *Bessie/bessie*[20] 'hypocoristic for Elisabeth/best friend', *rookie* 'e.g. a new recruit'.

[19] But this distinction is tricky.

[20] A random sample of the lexeme *Bessie/bessie* in *COHA* showed that the majority of the cases were the "embellished hypocoristic clipping" Bessie, and not *bessie* (for *best friend*).

The lexeme *Bessie* turned out to be one of the harder nuts to crack. Evidently, the capitalised form *Bessie* is an abbreviated form of the common name *Elisabeth*. It is difficult, however, because, as we can see, the suffix is attached to the truncated form (-*bess*-[21]), but the meaning is that of a typical hypocoristic, just like *Annie* and *Tommy*. Thus, the denotation is diminutive, but the formation process is that of truncation. With hesitation, it has been ruled non-diminutive. I argue it represents a hybrid form of the diminutive functions of truncation (see table 2.1 in section 2.2.2) and diminutive suffixation, and thus does not represent synthetic diminution only (base + diminutive suffix).

4.3 Diminutive frequency analysis

The discussion in this section will investigate the fundamental concept of diminutive productiveness in English and Spanish. Productivity has previously been defined as "the possibility of forming new words", and is the current definition applied in the analysis of diminutives in this chapter (Bauer 2001: 25, see 2.2.6). Thus, for a language to be considered productive on diminutive formation, diminutives must occur at a high rate in a language or particular area of language. The present analysis will look at language from two electronic corpora which contain language from both formal and informal areas of language as represented by the corpus genres of Spoken language, Fiction, News, Magazines and Academic texts (see 3.3.3 and 3.3.4 for a full account of the composition of the corpora). This study predicts that the most interesting findings will occur in the genres of Spoken language and Fiction (as these two categories contain the most informal language). The corpus data that will be looked into consist of 99 English diminutives and 84 Spanish diminutives.

The diminutive frequency analysis that takes place here will be two-part. The first part investigates the frequencies of the use of the lexemes relative to one another, both in terms of formation and relative frequency per lexeme, i.e. the English diminutives compared to the Spanish diminutives. The second part of the analysis looks at how productive the various diminutive suffixes within each language are in relation to each other (e.g. has *-ito* or *-illa* produced most diminutives, and which suffix demonstrates the highest number of relative frequency?). The first lexeme frequency analysis will answer the research question that is proposed here:

[21] It has been truncated both at the front and at the back.

Does English or Spanish demonstrate a higher relative frequency per diminutive lexeme?

The second part of this section, looking into productive suffixes, constitutes a much smaller part of the frequency analysis and less central. Throughout the discussion the emphasis will lie on the contrasts between English and Spanish diminutives, though occasionally, demonstrations of frequencies within English or Spanish will be highlighted to prove a point or if felt necessary. Corpus examples will be used to illustrate the frequent occurrence of a particular lexeme of suffix. Before the analysis can start, the two first steps of the method of contrastive analysis (see 3.2.1), i.e. description and juxtaposition of English and Spanish diminutives, will take place, to create a foundation for the frequency comparison.

4.3.1 Description

The two descriptions that will follow, one summarising Spanish diminutive formation, and the other summarising English diminutive formation, are based on "data-induced" research of the topic in both languages (see 3.2.1). This factor is most important in the contrastive analysis of two languages, so as to ascertain that the discussion is based on authentic linguistic contrasts rather than reflecting different methods used by linguists. In each of the descriptions, three variables will be discussed, namely (a) the repertoire of diminutive suffixes, (b) productive word classes of diminutive formation, and (c) the conditions for diminutive productivity. If needed, language-specific terminology will be applied to ensure the most accurate description possible for each language. Lastly, these two descriptions will be translated into a language-neutral account/description, to avoid confusion (see 3.2.1 for all the arguments).

Spanish diminution

(a) The full inventory of diminutive suffixes in Spanish, some suffixes more active than others, is conventionally estimated at a dozen suffixes, i.e. *-ito/a, -illo/a, -ico/a, -uelo/a, -ín/ina, -ino/a, -iño/a, -ajo/a, -ejo/a, -ete/a* and *-ijo/a*, which have been attested empirically[22] (e.g. Gooch 1967; Narváez 1970). The so-called group of emotives (diminutives, augmentatives and pejoratives) only accounts for a small portion of the several hundred

[22] Gooch (1967) and Narváez (1970) apply corpora of their own collections of diminutives. I have not come across Spanish research that treat diminutives through a quantitative corpus-based approach (e.g. through *CREA*, *CDE*, etc.).

derivational suffixes existing in Spanish. They are in return extremely productive; thereby the classification into emotive and non-emotive suffixes (Lang 1990: 91; 123).

(b) Spanish diminutive suffixation allows diminutive formation in most grammatical classes, i.e. nouns, adjectives, verbs, adverbs and pronouns. Gooch (1967: 33) observes that diminutives typically produce nouns from nouns and adjectives from adjectives, i.e. most often sticking to the same word class. Thus, as these are the two word classes of input (primitives), are described to be most productive in Spanish, we can then assume that the most productive diminutive outputs (derivatives) are nominal and adjectival diminutives (Normally the range of primitive bases are described in grammars of Spanish, not the other way around).

(c) How productive a language is in the formation of synthetic diminutives depends on the taxonomies of suffixal repertoires of the language, in addition to the flexibility and innovative qualities of formation (e.g. recursive application of suffixes, range of grammatical classes suffixes attach to). Spanish, as acknowledged by various research (e.g. Gooch 1967; Narváez 1970) applies most of these innovative diminutive functions, and can thus be characterised as extremely productive diminutive language.

English diminution

(a) English diminutive suffixation is generally characterised as an unproductive word-formation process, typically argued by the fact that this type of suffixes constitutes a miniscule part of the English derivational morphology, with only a small percentage of English suffixes having a diminutive meaning (cf. e.g. Quirk *et al*. 1972: 994; Katamba 2005: 59-63; Dressler and Barbaresi 1994: 84). Altogether, the inventory of diminutive suffixes in English consists of *-ie/-y/-ey*, *-ette*, *-let*, *-ling*, *-kin*, *-een*, *-en*, *-*(*e*)*rel*, *-et*, *-o*, *-a*, *-s*, *-er*, *-poo* and *-pegs*. However, only the first four are invariably characterised as productive today (Augustyn and Gniecka (2011: 34).

(b) As described by Bauer (1983: 221), English almost exclusively apply diminution in the grammatical class of nouns, i.e. producing nouns from nouns. The word class of adjectives only marginally applies diminution, whereas the process occurs even more sporadically with

verbs (Dressler and Barbaresi 1994: 112; Quirk *et al*. 1972: 994; Huddleston and Pullum 2002: 1678).

(c) Schneider (2003: 82) reports that research on synthetic diminutives has generally treated domains and discourse types that diminutives are not prone to occur in, as e.g. dictionaries. In addition, most diminutive research still relies on outdated assumptions about diminutive formation in English as foundations for the "research". It is only recently that diminutive studies have started to include other domains of language, and also applied corpora as data in the research. Augustyn and Gniecka (2011) empirically test the occurrence of diminutives in technical computer language and find that there are several productive diminutive patterns to be found (2011: 34). Research is still lacking in this linguistic field to be able to characterise diminution in English as productive or unproductive.

4.3.2 Juxtaposition

The common ground or *tertium comparationis* that has been established for the analysis of diminutive frequencies is the grammatical class of nouns, i.e. nominal diminutive derivatives. As previously discussed (see 3.2.1, also 4.3.1), this particular type of diminutive formation can be said to be productive in both English and Spanish. Let us take a look at what the corpora findings have to say about this.

4.3.3 Contrastive analysis proper and results

We can start looking at the data provided by the corpus searches in *COHA* and *CDE*. The first part of the analysis, as explained in the introduction of 4.3, will look at the most frequent diminutive lexemes in each language and compare their frequencies. Through this approach, an overview of the differences in productivity between English and Spanish diminutive formation should emerge, and we will see whether, indeed, Spanish is more productive on diminutive formation than English as previous research suggests. As previously discussed, it is hypothesised in this study, that:

> If it is the case that synthetic diminutives are more frequent in Spanish than in English, the present corpus findings from *COHA* and *CDE* should support this claim.

As such, the frequency investigation that is currently taking place will help expose the truth or falseness of this hypothesis. We will now look at some of the numbers from the corpus data analysis, where we want to find out whether English or Spanish demonstrates a higher relative frequency per diminutive lexeme. We will start by presenting the "Top Ten Frequency Ratings", which, as we can see from table 4.3, demonstrate a comparison of the most frequent English and Spanish diminutive formations in the time span 1800–2000.

Most interesting are the relative frequencies of the lexemes which are given per million words (pmw), as this suits the large sizes of the two corpora. The raw frequencies are only displayed here to give an impression of the size difference of *COHA* and *CDE* (the former being four times as big as the latter), and in addition, because these numbers have been used to calculate the relative frequencies.

Table 4.3: Top Ten Frequency Ratings of English and Spanish diminutive lexemes (c.1800–2000)

No.	Diminutive lexeme	Observed Frequency	Normalised Frequency
		Raw	PMW
1	PESETA(S)	3,234	78.2
2	SEÑORITO/A(S)	2,097	50.7
3	CIGARETTE(S)	17,968	44.2
4	BOLSILLO(S)	1,604	38.8
5	DARLING(S)	11,217	27.6
6	DADDY(IES)	10,525	25.9
7	CHIQUILLO/A(S)	801	19.4
8	PASILLO(S)	786	19.0
9	JUGETE(S)	743	18.0
10	CIGARRILLO(S)	743	18.0

As table 4.3 illustrates, the two most frequent diminutive formations are Spanish lexemes, i.e. the specialised diminutive *peseta(s)* 'peseta(s)', [23] demonstrating a relative frequency of 78.2 pmw, and secondly, *señorito/a(s)*, which has a relative frequency of 50.7 pmw. (Specialised diminutives, which are treated more thoroughly in the semantic analysis of 4.4, seldom have an expressive meaning, but rather a specialised sense of the primitive it derives from). The example of *pesetas*, here in (8), from the Spoken section of *CDE*, refers to the old Spanish currency unit:

(8) - Bueno... un divorcio, conforme a las tablas nuestras de honorarios en... de abogados, viene saliendo por unas ciento treinta mil *pesetas*, más o menos.
[Well.. a divorce, set our time tables in… of lawyers, comes out for some thirty thousand *pesetas*, more or less.]

(Spoken, 1900s, *CDE*)

The lexeme *señorito/a(s)* is similarly a specialised diminutive, at least historically. [24] The modern meaning of *señorito/a(s)* typically refers to a 'rich kid' or a '(degrading) young woman'. The corpus data demonstrate that the most frequent variant of *señorito/a(s)* appears in the form of feminine *señorita(s)*, as illustrated here in (9) and (10):

(9) - ¡¡¡Cómo, *señorita*!!! –exclamó la bruja, asumiendo una admirable actitud de sibila o pitonisa, y dando a su voz una inflexión severa -.¿Usted tiene un gato?
[How, *missy*!!! – the witch exclaimed, assuming an admirable prophetess-or-fortune-teller-like-attitude, and giving her voice a severe turning -.¿You have a cat?]

(Fiction, 1876, *CDE*)

(10) Lo de la... el programa de ayer de... de los chicos... es demasiado infantil, demasiado tonto, esos tres payasos que salen ahí; la *señorita* de las gafas cuadradas esas, la encuentro... vamos, son simpáticos ellos pero no, yo no sé si eso hará gracia a los niños.
[That thing… the show yesterday… for the kids… it's too childish, too stupid, those three clowns that are there; the *missy* with those square glasses, I find her… I mean, they are likable people but no, I don't see that this is amusing for kids]

(Spoken, 1900s, *CDE*)

[23] From *peso* 'coin' (Gooch 1967: 152). Remember that the diminutives in this study are from the period of 1800 to 2000, and thus logically *pesetas* is a much used word in this period as this was the main currency used in Spain and the Canary Islands before the advent of euros.

[24] The lexeme *señorito* used to refer to a title in the sense of 'young gentleman' (*WordReference* 2016).

Here, the diminutive expresses pejoration in both cases as in ‘missy; (degrading) young woman’. The sense of smallness is only felt vaguely; by choosing *señorita* (instead of e.g. *señora*) because the referents are young.

The most frequent English diminutive is in fact also a specialised diminutive (this function is not commonly acknowledged in descriptions of English diminutives). The English specialised diminutive here is *cigarette(s)* which expresses a relative frequency of 44.2 pmw. The relative frequency of *cigarette(s)* is very similar to *senorito/a(s)*, the second most frequent Spanish diminutive. This finding is interesting as it might indicate that if we include specialised diminutives in the class of diminutives, the productivity of the English diminutive category could be considered to be more productive than is generally perceived today (see table 4.3).

As a digression, it can be mentioned that the specialised diminutive *cigarette* is originally a diminutive of *cigar*, coming into the language as a borrowing from French for the sense of ‘a small cigar made of a little finely-cut tobacco rolled up in thin paper, tobacco-leaf, or maize-husk’ (*OED* 2016). Nowadays, however, the use of *cigarette* has become so widespread that we do not think about the diminutive sense it originally expressed. Thus, I argue that specialised diminutives, such as *cigarette*, should be included in the English diminutive category, as a new contribution to the existing repertoire of diminutive meanings in English, i.e. the senses of smallness, affection and pejoration, along the same lines as the Spanish diminutive category includes specialised diminutives. In Spanish this type has an important function in the language. As we will see in the semantic analysis of 4.4, there are further specialised diminutives to be found in English.

The remaining diminutive lexemes from the “Top Ten Frequency Ratings” presented in table 4.3 demonstrate the frequent occurrence of five Spanish diminutives in *CDE* and two frequent formations of English diminutives in *COHA*. All lexemes here demonstrate high relative frequencies per million words (for a diminutive, so frequent here means relative to other diminutive formations). Spanish *bolsillo(s)* ‘pocket’ from *bolso* ‘bag, handbag’ (38.8 pmw) is more frequent than the English formations *darling(s)* (27.6 pmw) and *daddy(ies)* (25.9 pmw). The two English hypocoristics do, however, beat the last four Spanish diminutives in relative frequency, i.e. *chiquillo/a(s)* ‘young boy/girl, kid, lad/lass’ (19.4 pmw), *pasillo(s)* ‘passage, corridor’ (19.0 pmw), *juguete(s)* ‘'toy(s); puppet, plaything (fig.)’

(18.0 pmw) and *cigarrillo(s)* 'small cigar' (18.0 pmw)[25]. Examples of the four Spanish lexemes can be seen in (10-13) and the two English lexemes in (14-17):

(10) En el liceo los *chiquillos* están fumando.
[In high school the *young kids* are smoking][26]

(Spoken, 1900s, *CDE*)

(11) Este año gastamos más de diez mil pesos en *juguetes*.
[This year we spent more than ten thousand kilos on *toys*]

(Spoken, 1900s, *CDE*)

(12) La señora Consuelo se levantó del sofá, sonrió a su madre como disculpándose y me hizo pasar a un pequeño escritorio junto a un *pasillo*.
[Mrs Consuelo got up from the sofa, smiled at her mother so as to say she was sorry and asked me to go over to a small desk by a *corridor*.]

(Spoken, 1900s, *CDE*)

(13) –Bien, entonces en las fiestas que tienen, se come. Se canta. ¿Qué más se hace? –Se toma. –Se toma, se canta. –Se baila. –Se baila, ¿qué más? –Se fuma. –Eh... creo que hay un cigarro, un *cigarrillo* muy famoso. ¿Cómo se llama? –Tabaco. –Ah... el tabaco. –Yo he oído hablar de otro nombre... y tiene el nombre de la capital, de la ciudad... del país de ustedes... ¿Cómo se llama? Puede que esté equivocado... ¿el habano?
[–I mean, in the parties they throw, you eat. –You sing. What more do you do? –You drink. –You drink, you sing. –You dance. You dance, what more? You smoke. –Eh… I think there is a cigar, a very famous *cigarette*. What's it called? –Tabaco. Ah, the tabaco. –I have heard another name… and it is the name of the capital, of the city… of the country you're from… What is it called? I might be confused… the Havannan?]

(Spoken, 1900s, *CDE*)

[25] The meanings of the Spanish lexemes are taken from Goosh (1967: 8; 83; 90; 91; 93; 104) and *WordReference* (2016).

[26] I have rendered *los chiquillos* by *the young kids* in the English translation of this Spanish diminutive. This type of construction with *young* is an analytic diminutive construction, which normally is expressed periphrastically through *little* (or young, dear), instead of the synthetic way of adding a suffix to the base. Possible synthetic translations are *the kiddoes* or *the kiddies* for *los chiquillos*, but they felt a little unnatural in this context. Spanish *chiquillos* could also simply be translated as *children*. The three possibilities of renderings into English of this lexeme demonstrate the width of use (diminutive) suffixes have in the Spanish language, as suffixation is more common than e.g. compounding, which is the main English word formation process.

Evidently, all the Spanish diminutives here occur in the informal language of everyday conversations, TV-shows and fiction. In contrast, the English diminutives are found in fictional texts, i.e. *daddy(ies)* and *darling(s)*. The explanation of this difference lies, however, in the absence of a Spoken genre in the English corpus. *COHA* only consists of the genres of Fiction, Magazine, Newspaper and Non-fiction (see 3.3.3). Sentences (14-17) illustrate this fact:

(14) What matters is that families love each other. I know you love your *daddy*. " Yep", Brett squealed, jumped up. " I love you, *Daddy*. " # "I love you, too."

(Fiction, 2009, *COHA*)

(15) She appeared at the top of the staircase with Mama hovering behind her, crowing, " Here she is! " as *Daddy* and Rand walked in from the sitting room.

(Fiction, 1995, *COHA*)

(16) Now listen carefully, *darling*. Fill the tank of my fastest car. Pack something for yourself. Only what's necessary. Don't ask questions. Just do as I tell you.

(Fiction, 1979, *COHA*)

(17) Michael turned to her, still sobbing, in a state of terror. Cass held him. "Come on, *darling*, everything's all right, hush now, darling, come on".

(Fiction, 1962, *COHA*)

With table 4.3 as the point of departure, this frequency comparison has provided a initial overview of the frequency trends of English and Spanish diminutives the last 200 years (c.1800–2000). Spanish has come out strongest, declaring itself as the more productive of the two, i.e. in terms of the number of productive formations produced in this period (seven diminutives included in the Top Ten Frequency Ratings). English demonstrates fewer productive formations than Spanish (three diminutives in the Top Ten Frequency Ratings). However, the three English diminutives (*cigarette(s), darling(s), daddy(ies)*) that are represented among the top ten are, on the other hand, quite frequent per million words (relative to the other lexemes included in the top ten), beating four Spanish formations (*chiquillo/a(s), pasillo(s), juguete(s), cigarillo(s)*). Also, as a side note, as this primarily is a frequency comparison (the semantic comparison will follow in 4.4), the two forms *daddy(s)* and *darling(s)* are the two forms that express the clearest sense of diminution, i.e. smallness,

which is the fundamental meaning of diminutives (see e.g. Jurafsky 1996: 534 or 2.2.1). Spanish, in contrast, only has one fundamental diminutive represented among the top ten, namely *señorita*.[27] That said, we will now carry on to look at English and Spanish separately to determine the most productive diminutive suffixes in each diminutive language.

Most productive diminutive suffixes

The tables (tables 4.4 and 4.5) that will be discussed in this section demonstrate the most frequent diminutives per language sorted by relative frequency, one for Spanish and one for English. This perspective provides a larger insight into each language, as the ten, instead of five, most frequent diminutives will be given in each table. As table 4.4 shows, the most productive Spanish diminutive suffix, in terms of the number of formations produced (in the last two hundred years), is the suffix *-illo/a(s).*The suffix has produced as much as five diminutives among the Spanish Top Ten and thus stands out as the most productive diminutive suffix in Spanish by far. We have to be aware of the fact that these tables only represent the very tip of the iceberg of the state of Spanish diminutive suffixation. As such, table 4.4, and table 4.5 which will follow swiftly, can only be considered a snapshot of productive diminutive patterns in each language. The suffix *-illo* has produced the diminutives *bolsillo(s), pasillo(s), cigarillo(s)* and *guerrilla(s)*, which are cases of specialised diminutives, in addition to c*hiquillo(s)*, which is a diminutive-pejorative formation. This finding shows that the suffix *-illo/a* has a strong tendency to produce diminutives with a specialisation of meaning.

[27] The corpus hits for *señorito/a(s)* shows that *señorita* is by far the most frequently used form of the four (masculine singular, feminine singular, masculine plural and feminine plural), which is why I only put the feminine singular form in this little summary, as it best represents the use of this lexeme.

Table 4.4: Spanish Top Ten (1800-2000) from *CDE*

Nr.	Diminutive lexeme	Frequency	
		Raw	PMW
1	PESETA(S)	3,234	78.2
2	SEÑORITO/A(S)	2,097	50.7
3	BOLSILLO(S)	1,604	38.8
4	CHIQUILLO/A(S)	801	19.4
5	PASILLO(S)	786	19.0
6	JUGETE(S)	743	18.0
7	CIGARRILLO(S)	743	18.0
8	GUERRILLA(S)	588	14.2
9	MESETA(S)	462	11.2
10	BANQUETE(S)	455	11.0

After *-illo/a*, the Spanish Top Ten "snapshot" representation shows that the suffix *-ete/a* has been second most productive in the formation of diminutives in this period. The diminutives produced by *-ete/a* are *juguete(s)* 'toy' (from *juego* 'game'), *meseta(s)* 'tableland, plateau' (from mesa 'table'), and *banquete(s)* 'banquet' (from *banco* 'bench') , which precisely like the suffix *-illo/a,* are instances of specialised diminutives. Only *juguete(s)* is ambiguous here, being able to express an additional diminutive-pejorative sense, as in 'puppet, plaything (fig)'. The example of *juguete* in (11), as previously demonstrated, shows the use of *juguete* in the specialised sense of 'toy'. Unfortunately, as this study applies a primarily quantitative approach to diminutive use, the discussion of the lexemes is not comprehensive enough to include an analysis of how the lexemes are used in the corpora. It would have been interesting to see in which meaning e.g. *juguete* is most used in.

As table 4.4 reveals, the third most productive diminutive suffix in Spanish is *-ito/a*. The Spanish Top Ten demonstration only shows one instance of an *-ito/a*-formation, i.e. in the form of *señorito/a(s)*. This is somewhat surprising, as *-ito/a* is the most elementary and basic diminutive suffix there is in Spanish. To its defence, this one formation, *señorito/a(s)*, is in return the second most productive of the top ten formations per million words. A possible explanation for this distribution is that specialised diminutives do not require an informal

setting to occur in language, which is the case for most other diminutive senses, e.g. the diminutive senses of smallness, hypocorism and pejoration.

Moving on to table 4.5 and the English Top Ten diminutive formations, we find that, in the same period of time (c.1800s-2000s), the suffix *-ie/-y/-ey* has been the most productive English diminutive suffix. This suffix is commonly known to signal affection (see 2.2.6), as can be deduced from the formations of the suffix *-ie/-y/-ey* in table 4.5:

Table 4.5: English Top Ten (1810-2009) from *COHA*

Nr.	Diminutive lexeme	Frequency Measurings	
		Raw	PMW
1	CIGARETTE(S)	17,968	44.2
2	DARLING(S)	11,217	27.6
3	DADDY(IES)	10,525	25.9
4	PONY(IES)[28]	6,226	15.3
5	PAMPHLET(S)	3,859	9.5
6	COOKIE(S)	3,055	7.5
7	MAMMY(IES)|MOMMY(IES)|MUMMY|MAMY(IES)	2,908	7.2
8	BUDDY(IES)	2,813	6.9
9	BUNNY(IES)	2,789	6.9
10	BUGGY(IES)	2,674	6.6

In total, seven formations of the suffix *-ie/-y/-ey* can be found in the English Top Ten diminutive demonstration, i.e. *daddy(ies)*, *pony(ies)*, *cookie(s)*, *mammy(ies)*[29], *buddy(ies)*, *bunny(ies)* and *buggy(ies)*. Three of these diminutives are strictly hypocoristics (*daddy*, *mammy*, *bunny*), while the rest are ambiguous (*pony*, *cookie*, *buddy* and *buggy*). Diminutive *pony* can refer to 'a smaller horse of any breed' (smallness), whose use is not common any more, as this meaning has become specialised, denoting now 'a special type of horse breed'

[28] Probably a borrowing from French (*OED* 2016)

[29] And the variations of *mammy(ies)*, i.e. *mommy(ies)*, *mummy*, and *mamy(ies)*.

(specialisation), in addition to the regional sense of ‘a small glass or measure of alcohol’ (specialisation) (*OED* 2016). Diminutive *cookie* can refer to ‘a small flat sweet cake’ (smallness) or, for instance, ‘an attractive young woman’ (specialisation).[30] The lexeme b*uddy* is either used as in an affectionate application as in ‘a good friend’ (hypocoristic), or a specialised sense, as e.g. ‘a member of a pairing, esp. at a school or workplace’ (specialisation). Lastly, the lexeme *buggy* is variously used to express the denotation of ‘a little bug’ (smallness) or ‘a cute little bug’ (hypocoristic) or e.g. ‘a special type of car’ (specialisation)[31].

The other three suffixes represented in the English Top Ten demonstration of table 4.4 are *-ette*, *-ling* and *-let*. Each of these three suffixes is represented by one diminutive formation each, i.e. *cigarette(s), darling(s)* and *pamphlet(s).* The lexemes *cigarette* and *darling(s),* have previously been discussed; the former, *cigarette,* being a case of specialised diminutive, and the latter, *darling*, a hypocoristic formation and common term of endearment, as in ‘a person who is very dear to another; the object of a person's love; one dearly loved’. Diminutive, *pamphlet(s),* has not previously received attention. This diminutive denotes smallness, referring to a smaller version of something larger, i.e. ‘a short printed work of several pages fastened together without a hard cover; a booklet; a leaflet’.

4.3.4 Frequency trends

To sum up, the comparison of the most productive diminutive lexemes in English and Spanish in the last two hundred years shows that Spanish is the most productive diminutive language of the two, the corpus data (as presented in tables 4.3, 4.4 and 4.5) demonstrating a higher number of frequent Spanish diminutive formations making the Top Ten Frequency Ratings (seven of ten) (see table 4.3). English has fewer diminutive formations represented among the Top Ten Frequency Ratings (three of ten), but does not show signs of unproductiveness, rather the contrary, as the diminutives that *are* represented among the top ten are quite frequent per million words (relative to the other seven diminutives).

In terms of frequent denotations expressed, the category of specialised diminutives is the clear winner on the Spanish side, having produced five of seven specialised diminutives among the Top Ten Frequency Ratings (in English and Spanish, see table 4.3). The Spanish

[30] There are several specialised meanings of the lexeme *cookie* (see *OED* 2016).

[31] Note: The uses of these ambiguous lexemes are not extensive.

Top Ten demonstration similarly shows a clear dominance of specialised diminutives, this category making out eight of ten clear cases of specialised diminutives (*señorito/a(s)* is not included here, as its use in the *COHA* predominantly refers to the diminutive-pejorative use of the lexeme, the specialised sense of *señorito* largely being historical). The clear winner in English, on the other hand, is the hypocoristics category, being represented by two of three lexemes in the Top Ten Frequency Ratings (see table 4.3) and six of ten instances in the English Top Ten demonstration (see table 4.5).

These "snapshot" perspectives of the most frequent productions in English and Spanish demonstrate that the English language is best represented in one of the fundamental diminutive denotations, i.e. the hypocoristic category (e.g. *daddy*, *darling*, *bunny*). Interestingly, this means that English is actually more frequent in the basic diminutive sense of hypocorism than Spanish (see table 4.3). We do, however, need to include a larger list of lexemes to see if this trend continues (the discussion of semantic categories of 4.4 will look at the larger picture).

Furthermore, the most productive diminutive suffix in Spanish is *-illo/a*, which has a strong tendency to produce diminutives with a specialisation of meaning. In English, the most productive diminutive suffix is the affectionate marker *-ie/-ey/-y*, producing a mix of hypocoristics, specialised diminutives and diminutives expressing smallness.

Interestingly, also, is the finding of two specialised diminutives (*cigarette* and *pony*) among the English data, which represent a diminutive type which previously has not been acknowledged by English research. This study argues that specialised diminutives should be endorsed as a diminutive category, as a new contribution to the existing repertoire of diminutive meanings in English, i.e. the senses of smallness, affection and pejoration, along the same lines as the Spanish diminutive category includes specialised diminutives (in Spanish this type has an important function in the language). Consequently, it is still too early to see if there is a real pattern of productive specialised diminutives in English, but this study will regardless argue that this diminutive type exists in English. Furthermore, such a finding could indicate that if we include specialised diminutives in the class of diminutives, the productivity of the English diminutive category could be considered to be higher than is generally perceived today (see table 4.3).

4.4 Semantic distribution analysis

This section will investigate another fundamental feature of diminution, i.e. the distribution of diminutive meanings in English and Spanish. This analysis will through a surface semantic classification look at the use of the most frequent diminutive lexemes the two languages and classify them into five diminutive categories. This classification of diminutives into five different categories will hopefully be able to answer the second research question posed here:

> In terms of semantic distribution, are the English or the Spanish diminutives more innovative?

The corpus data that will be analysed are the same 99 English diminutives and 84 Spanish diminutives that were applied in the lexeme frequency analysis of 4.3. Logically, as these lexemes are the same as in the previous discussion, they are all nominal diminutives. Thus, the analysis of diminutives in the discussion here will be able to say something about the semantic use of nominal diminutives in English and Spanish. As previously discussed (see 4.1.1), this semantic analysis will, like the frequency analysis of 4.3, contribute to either confirm or reject the hypothesis of this study:

> If it is the case that synthetic diminutives are more frequent in Spanish than in English, the present corpus findings from *COHA* and *CDE* should support this claim.

The method applied in this semantic distributional analysis of English and Spanish diminutives is the three-step approach of contrastive analysis (see Lefer 2011: 650). The three steps of analysis include a description, juxtaposition and an analysis proper. In the first stage, I will describe the conventional diminutive meanings of English and Spanish separately. The second stage will compare these meanings and establish a *tertium comparationis* (TS) for the comparison (cf. 3.2.1). The final stage is the contrastive analysis proper, which will present, analyse and discuss the corpus findings, and point to differences and/or similarities of the languages. This last stage constitutes the main part of this analysis in 4.4.

Methods that will be used to demonstrate the differences between the two languages' diminutive patterns are, first, the relative frequencies of the diminutive categories, and second, a few statistical methods (mainly log likelihood), which will be used to test the

significance of the differences found. The statistical methods applied will be addressed in the preliminary section of the analysis of 4.4 (see 4.4.1). In this section, a few other questions concerning the focus applied in this analysis will also be addressed. The remaining build-up of the analysis is as follows: a description of the two languages in terms of meaning will be provided, which thus in the juxtaposition stage will establish a TC for the comparison. Here, an account of the five semantic categories applied in the semantic distributional analysis will be given. In addition, the reason for including these five categories will be stated. Furthermore, the analysis proper will take place. The analysis will close by discussing a few semantic trends indicated by the analysis.

4.4.1 Preliminary questions

Before setting out to analyse the corpus data form *CDE* and *COHA*, and present the findings from the semantic distribution analysis, a few issues must first be addressed. First, a brief introduction to the quantitative, but also partly qualitative, approach applied in the analysis will be stated. This research approach will be discussed in relation to the ambiguity that many diminutives express, and the advantages and disadvantages of a mainly qualitative approach. Second, the main statistical method applied in this analysis will be discussed, i.e. the method of log likelihood. Here, the usefulness of statistical methods in this type of research will be argued for.

Quantitative focus and ambiguity

The present study mainly applies a quantitative approach to the topic of diminutive suffixation in English and Spanish, both in respect to the frequency analysis of 4.3 and the current analysis of semantic distributional analysis. As this semantic analysis aims to provide a surface mapping of the distribution of semantic meanings in English and Spanish, which can be used to say something in general about differences between English and Spanish diminutive use, a quantitative approach to the topic is the appropriate choice.

However, the study also concerns discussions of ambiguous diminutive cases and looks at meanings expressed by the corpus data. Thus, these actions require some qualitative investigations into the corpus material. The qualitative investigations will consist of some in-depth analyses of samples of the hits for a given lexeme. This approach is mainly applied in the case of ambiguous diminutives to discover which meanings are mostly used in the corpus data (for example if the dictionary lists three possible meanings for the lexeme). For instance,

some lexemes express three meanings (as with e.g. *semilla*, see 4.2.4), where only one of them is diminutive. To determine if the lexeme is used in a diminutive way in the corpus a random sample investigation is this carried out.

If, contrastingly, a qualitative approach had been applied in a diminutive study, this would enable a more thorough and in-depth analysis of diminutive conduct, where the lexeme would be the centre of attention. Such a study would, however, have to restrict its scope to include much fewer lexemes than the present study looks at.

The disadvantage of the mainly quantitative approach applied in the present study is, for instance, that the actual meanings represented by each lexeme in the corpus cannot be studied in detail. The present study must, instead, content itself with seeking aid in dictionaries to account for the meanings expressed by the diminutives attested in the corpora. Unfortunately, dictionaries do not always provide the full extent of meanings possible for a given diminutive lexeme. Neither can a dictionary guarantee that the meaning represented by a lexeme in the corpus represents any one, or any at all, of those listed in the dictionary. This study therefore does not pretend to provide all the correct meaning(s) for the diminutives discussed in the study, but will attempt to check meanings that are particularly ambiguous. A suggestion for further study is therefore to carry out a similar investigation as the present one, but from a qualitative approach, in order to correct any of my mistakes, and to confirm or reject the semantic trends of diminutive use claimed in the present study.

The method of log likelihood (LL) applied in this study

This section will principally discuss the statistical method of log likelihood (LL) which will be applied in the analysis proper of this semantic distributional analysis. In addition, a short explanation of the marginally applied statistical methods of *p*-value and the Bayes factor tests will be provided.

The log likelihood (LL) test is a much used method in the branch of inferential statistics. Typically, the LL test operates as a means to test a hypothesis, to discover whether observed differences between two datasets are statistically significant. The *p*-value is similarly calculated to test the statistical significance of the LL, in order to determine whether or not the LL score is statistically significant, i.e. if it is due to chance or not. For instance, a *p*-value of p=0.03 shows that it is 97% certain that the LL is not due to chance. Both the LL method and the *p*-value are applied to demonstrate "how confident one can be that the observed relationships or differences are not due to chance alone" (Levshina 2015: 129). This

means that if the observed differences are not coincidental, one can logically conclude that they are statistically significant. In the comparison of unevenly sized datasets, like my own (*COHA* being four times larger than *CDE*), such demonstrations are particularly useful.

In the present analysis, the LL test aims to discover whether the differences in use between a given semantic category in English compared to Spanish are statistically significant. Normalised frequencies do not demonstrate whether the difference in use is significant; they only demonstrate that there is difference, and thus, do not alone offer a sufficient demonstration of the differences of use between two corpora.

Earlier applications of the LL method in research include the comparisons of vocabulary according to social factors (age, gender, etc.) (cf. Rayson *et al*. 1997) and native and non-native lexical variation in writing revealing e.g. major word-classes used (cf. Granger and Rayson 1998), to name but a few. A more recent study, more similar to my own, is Rayson and Garside (2000), which applies the method of log likelihood to compare the differences of semantic categories in a systems engineering domain from one corpus to another (2000: 3).

The LL test is often regarded as the ideal test for the measuring of statistical significance. Importantly, in relation to size, it does not assume that two sets of frequencies are already normalised or comparable, but takes the corpora sizes into consideration, measuring the relative frequencies. Similarly, as described by Rayson and Garside (2000: 2), when dealing with high frequency values, the LL test is the best-suited choice, as it will be less prone to overestimate the significance from the oversized (numbers of the one dataset) (2000: 2). In contrast, the chi-square test, a much-used method prior to the advent of LL, only compares the observed values and expected values, disregarding the normalised frequencies of the datasets. The problem then occurs if "the expected frequency is less than 5", resulting in an untrustworthy chi-square value (Rayson & Garside 2000: 2).

To measure the log likelihood, two frequency lists, one for each corpus, must be produced. Typically, one makes a frequency list of words, as in my case, but it can also be a "part-of speech (POS) or semantic tag frequency list" (Rayson & Garside 2000: 2). Then, one enters four numbers into a LL calculator[32], namely the observed values one wants to compare and the total sizes of the corpora. The number that is then calculated as the log likelihood can be any number >0. For the difference to be considered statistically significant, the likelihood

[32] The calculator used in my analysis: http://ucrel.lancs.ac.uk/llwizard.html

score, must be greater than 3.84 (Oakes 1998: 266). The larger the score is, the more significant the results are, and the less likely it is that the observed relationship is due to chance (Levshina 2015: 129). Similarly, a probability value *p* of 0 demonstrates that the difference is highly significant statistically, whereas if the *p* value is near to 1, the difference is most likely due to chance. The hypothesis can only be acknowledged when the level of significance, the *p*-value, is less than 0.05 (i.e. $p<0.05$, or more than 95% confident).

Lastly, to control that the LL score and the *p*-value are correct, one can apply a test that measures the effect size of a difference. The main method applied to test the effect size of LL is the Bayes factor. It is vital to distinguish between the effect size, on the one hand, and statistical significance and LL on the other. Effect size demonstrates, as stated by Levshina (2015: 129), "the strength of a relationship or the magnitude of a difference". However, a strong effect size does not automatically imply significance, and vice versa (Levshina 2015: 129). Statistical significance or LL, on the other hand, as previously discussed, only demonstrates how certain one can be that the observed difference is not due to chance (Levshina 2015: 129).

4.4.2 Description

This section will provide a description each of English and Spanish diminutive meanings to ensure that like is compared with like (e.g. that the use the smallness sense of English diminutives has roughly the same application in Spanish). This will then provide a sound basis for the comparison, which will help to establish a TC for the semantic comparison in this analysis.

Spanish diminution

It is important to remember that Spanish, in contrast to English, applies derivation as the main word formation device, while English prefers compounding in the production of new lexis. One result of that is that Spanish also uses diminutive suffixes to produce diminutives with a specialised meaning. This use of diminutive suffixes differs from the normal function of diminutives, which is to alter the meaning of the primitive emotionally. Specialised diminutives are thus also similar to lexemes produced by non-emotive suffixes.
The range of meanings attested for Spanish diminutives that the background research for this study has come across is: smallness, hypocorism (i.e. affection, closeness, familiarity), pejoration, specialisation of meaning, attenuation, intensification, persuasion, courtesy,

respect, adding of colour, and even augmentation (the opposite of diminution) (Gooch 1967: 7-8). The range of meanings listed here is expressed through a large repertoire of diminutive suffixes, which all have their own distinctive shade of meaning (Lang 1990: 102; *Nueva gramática de la lengua española* 2010: 166).

English diminution

English diminutive meanings are, in contrast, considerably more restricted. Three meanings can be said to cover the range of meanings expressed by English synthetic diminutives: smallness, hypocorism (i.e. affection, familiarity, closeness) and pejoration. A function which this study argues should be included in the English diminutive category is the type of specialised diminutives that is used extensively in Spanish. This type of diminutives has previously not been acknowledged by diminutive research, but will nevertheless be applied in the discussion of semantic meanings.

The repertoire of diminutive suffixes in English is rather small, consisting of the suffixes *-ie/-y/-ey*, *-let*, *-ling* and *-ette*. There are also number of other diminutive suffixes which receive little attention in this study, but which have been included in other diminutive studies (this is not the full list), i.e. *-kin, -o, -s -a, -et, -(e)rel, -er, -een* and *-en*. Some of these can be said to still be productive (e.g. *-o,* as in *kiddo*), while others are historical or largely unproductive.

4.4.3 Juxtaposition

The two separate descriptions of the various denotations of the diminutive category in English and Spanish demonstrate that there are three senses which are similarly applied in both languages, i.e. the expressions of smallness, hypocorism (affection, familiarity, closeness) and pejoration[33]. Thus, these three meanings will constitute the *tertium comparationis* (TS) for the semantic diminutive analysis here. In addition, I will also include specialised diminutives in the TC of this study, as they constitute an essential function of Spanish diminutives. I argue that specialised diminutives also exist in English, despite the lack of mention of it in English research. I would logically assume that the lack of acknowledgement of this function stems

[33] It may be noted that Spanish diminutives have a much more nuanced set of senses, in which a diminutive expressing smallness can be e.g. (a) generally affectionate, (b) frequently affectionate or some other favourable implication, or (c) often affectionate, but also used ironically with augmentative-pejorative force, etc. The division into categories of smallness, pejoration, etc., will include all such nuances in a broader category of e.g. ‘smallness’.

largely from the approaches applied in English diminutive research to date. Based on the previous research I have reviewed, the most comprehensive study of English diminutives so far is Schneider (2003), which looks at the pragmatic functions of the diminutive, but does not go into detail about the denotational and/or connotative meanings of diminutives.

Ultimately, as both English and Spanish diminutives demonstrate the ability to express ambiguity, I will include an ambiguous category to demonstrate this innate nature of diminutives. The present study is aware of the fact that ambiguity is not a meaning in itself, but this analysis will, however, apply it as a "semantic category" in the discussion, to demonstrate the ability of diminutives to express various senses and the extent to which this ambiguous use occurs. To sum up, these five diminutive senses will constitute the TS of this semantic diminutive analysis.

The five semantic categories

A brief description of the diminutives included in the various categories will be accounted for here.

Smallness

Diminutives included in the category of smallness are lexemes such as *chiquillo* 'young boy', *casita* 'small house', *mesita* 'small table', *pajarito* 'small bird', *pamphlet*, *puppy*, etc. The meaning essentially expressed by the lexemes in this category is smallness of any kind, e.g. size, degree. In addition, these diminutives express a variety of sub-meanings, e.g. implications of pejoration, irony, affection or other favourable impressions (extensively so for both languages). Based on the observation of lexemes expressing smallness, in addition to a second connotative meaning, one can with certainty say that all diminutives expressing smallness are also ambiguous (see diminutives analysed as smallness-diminutives in Appendix A).

Hypocoristics

The primary function of hypocoristics is to express affection, closeness, cuteness and/or familiarity, as in the common applications of *daddy*, *granny*, *Jimmy*, *hijita* 'little girl', *ratito* 'a short while', *casita* '(delightful) little house', *mesita* '(cute) little table' and *pajarito* '(sweet) little bird'. The meaning expressed by hypocoristics is, however, almost always combined with the element of smallness. Proper names often acquire hypocoristic forms, as people like to assign nicknames for people they have a more intimate relationship with.

Pejoration

The pejorative category comprises of diminutives which invariably express a combination of smallness and pejoration, i.e. negatively loaded diminutives expressing negative implications such as contempt and derogation. The lexemes included in this category are always ambiguous. In Spanish, diminutive-pejorative lexemes have a tendency to acquire specialised senses, as e.g. *juguete*, which can refer to the pejorative sense of a 'plaything' or the specialised sense of a 'toy'.

Specialisation

Specialised diminutives have an important function in the Spanish language. This type of diminutives is not expressive in manner, i.e. expressing cuteness or another diminutive emotion. Instead, these lexemes have acquired a change or specialisation of meaning through the addition of a diminutive suffix, which can be altogether different from the meaning of the primitive, as illustrated with instances from the Spanish corpus, e.g. *peseta* 'peseta' from *peso* 'weight (coin)', *bolsillo* 'pocket' from *bolso* 'bag, handbag', *pasillo* 'passage, corridor' from *paso* 'step' (Gooch 1967: 90; 93; 152). Some specialised diminutives denote an alternative meaning of smallness or pejoration, as illustrated by *camarín* from *cámara* 'chamber', which can refer to the changed sense of 'dressing room' or the more diminutive-pejorative sense of 'small chapel' (Gooch 1967: 75; *DRAE* 2016).

In current accounts of English diminutives (at least that this study has come across), there is no mention of specialised diminutives. For this reason, this study does not propose that this type of diminutives has an essential function in the English language. This study does, however, argue that this type of diminutives exists in English, but has not been recognised as such. The corpus data studied in the present investigation implies as much, as can be illustrated by the following findings from the English corpus, i.e. *buddy* 'a pair of miners' or alternatively 'member of a pairing at a school/workplace' from *bud*/*brother*, *caddie* 'a member of a corps of commissionaires in Edinburgh in the 18th' from French *cadet* 'a son or brother; a gentleman who entered the army to learn the profession' (*OED* 2016). Additional examples are *cigarette, buggy* and *tablet.* (See Appendix A for all findings of specialised diminutives).

Ambiguous category

The ambiguous category comprises all diminutives expressing more than one meaning. Thus,

the lexemes included in the ambiguous category are also included in one or several of the other four semantic categories applied in the present semantic classification.

4.4.4 Contrastive analysis proper and results

We will now start the investigation of the corpus data and present the results of the semantic distribution analysis that has been undertaken. As previously stated, the corpus data consists of 99 English diminutives and 84 Spanish diminutives which have been analysed according to their meaning(s), and categorised into five different groups. The contrastive analysis proper that takes place in this section will attempt to shed some light on the differences between English and Spanish in terms of distribution of diminutive meanings as it has been manifested in the last two hundred years.

This section presents the data from the semantic diminutive analysis of English and Spanish lexemes, based on two methods: a normalised frequency comparison and a statistical significance test (principally a log likelihood test) of the same data. The use of these two methods will be applied in two separate discussions, which each will look at the semantic differences between diminutives in English and Spanish. The first analysis will focus on the relative frequencies of each semantic category, which will be applied as a means to contrast the productiveness of the various categories in English and Spanish (e.g. how frequent the application of pejorative-diminutive meanings is in English compared to Spanish). Secondly, the differences will in the next discussion be tested for statistical significance through the methods of log likelihood, *p*-value and Bayes factor (see 4.4.1).

Presentation of the corpus data

The complete classification of lexemes into semantic categories can be found in Appendix A, but was found too large to be a part of the main body of this study (I would encourage the reader to take a look at the table in Appendix A to get an impression of what the frequencies of tables in 4.6, 4.7, 4.8 and 4.9 represent). Table 4.6 demonstrates the division of lexemes into semantic categories in Spanish. Table 4.7 demonstrates the same for the English data. The numbers given in the tables are the number of diminutive formations found in the two respective corpora, i.e. only the most productive diminutives from each language, in addition to the percentages for each diminutive category based on the raw numbers of lexemes. Note that the ambiguous category is presented last in the table, after the 'Total'-column, as the lexemes classified into this group are also represented in at least one of the other diminutive

categories. Therefore, ambiguous formations should be excluded from the total sums of lexemes.

Table 4.6: Semantic classification of Spanish diminutive lexemes [34]

	Smallness	Hypocorism	Pejoration	Specialisation	Total	Ambiguous
No.	14	8	13	49	84	14
Percentage	16.7%	9.5%	15.5%	58.3%	100.0%	16.7%

Table 4.7: Semantic classification of English diminutive lexemes

	Smallness	Hypocorism	Pejoration	Specialisation	Total	Ambiguous cases
No.	27	27	19	26	99	29
Percentage	26.3%	27.3%	27.3%	19.2%	100.0%	14.1%

These two demonstrations appropriately describe a first overview of the productivity of formation within each of the semantic categories of smallness, hypocorism, pejoration, specialisation and ambiguous cases within Spanish and English, respectively. However, due to the great difference in size between the two corpora (*COHA* is four times larger than *CDE*), the numbers presented in tables 4.6 and 4.7 may misrepresent the relative difference between the two languages. As such, the present demonstration says little about the productivity of diminutive formation between the two languages relative to one another. The fact that there are 99 diminutive formations found in the English data, while the Spanish data only is represented by 84 diminutives should ring a bell. To demonstrate some of the diminutive formations found in the corpus data, a few lexemes from each diminutive category are presented in tables 4.8 and 4.9:

[34] *The sums given here are the raw numbers of lexemes analysed to have that particular semantic meaning in the Spanish data

Table 4.8: Examples of diminutive findings from *CDE*[35]

Smallness	Hypocorism	Pejoration	Specialisation	Ambiguous cases
señorito/a(s)	*casita(s)*	*señorito/a(s)*	*peseta(s)*	*señorito/a(s)*
'e.g. young woman'	'(delightful)	'missy'	'peseta'	*chiquillo/a(s)*
chiquillo/a(s)	little house'	*chiquillo/a(s)*	*pasillo(s)*	*juguete(s)*
'e.g. young boy'	*mesita(s)*	'e.g. kids'	'e.g. corridor'	casita(s)
casita(s)	'cute little	*juguete(s)*	*cigarrillo(s)*	*botín(es)*
'small house'	table'	'plaything'	'small cigar'	*mesita(s)*
botín(es)	*pajarito(s)*	pajarito(s)	*juguete(s)*	
mesita(s)	'(sweet) little	'"dicky" bird'[36]	'toy'	
'small table'	bird'	pajarillo(s)	*festín(es)*	
pajarito(s)	hijita(s)	'(miserable)	'feast'	
'small bird'	'little girl'	little bird'	*botín(es)*	
	botín(es)	*marica(s)*	'e.g. ankle	
	'(baby's)	'pansy, sissy'	boots'	
	bootie'			

[35] The corresponding meanings of these diminutive findings are in the majority of the cases taken from Gooch (1967). As such, there are many other denotations possible for these diminutives, and Gooch's definitions (1967) are only one solution.

[36] This pejorative sense is perhaps equivalent to Norw. 'forbanna gjøk'

Table 4.9: Examples of diminutive findings from *COHA*

Smallness	Hypocorism	Pejoration	Specialisation	Ambiguous cases
cigarette(s)	*darling(s)*	*missy(es)*	*cigarette(s)*	*cigarette(s)*
pony(ies)	*daddy(ies)*	*puppy(ies)*	*cookie(s)*	*cookie(s)*
pamphlet(s)	*buddy(ies)*	*bully(ies)*	*buddy(ies)*	*buddy(ies)*
cookie(s)	*bunny(ies)*	*pussy(ies)*	*buggy(ies)*	*buggy(ies)*
buggy(ies)	*granny(ies)*	*sapling(s)*	*tablet(s)*	*tablet(s)*
tablet(s)	*puppy(ies)*	*nanny(ies)*	*brownie(s)*	*puppy(ies)*
puppy(ies)	*auntie(s)*	*panty(ies)*	*sapling(s)*	*missy(es)*
hamlet(s)	*missy(es)*	*weakling(s)*	*sweetie(s)*	
	bully(ies)	*underling(s)*	*kindling(s)*	

The denotations added for the Spanish diminutives are mostly senses Gooch (1967) has ascribed to these diminutives. This study does not claim that the same meanings are reflected in the corpus examples. That would require a thorough qualitative analysis of all hits for each of lexemes 99 Spanish lexemes and 84 English lexemes, which is beyond the scope of this study.

As regards the previously discussed "inadequacy" of tables 4.6 and 4.7, a more detailed table will now be presented. Table 4.10 importantly illustrates the relative productivity per diminutive category between English and Spanish. The demonstration here, in table 4.10, includes a more detailed overview of the formation of diminutives in English and Spanish, the table providing both observed frequencies and relative frequencies. Hence, this demonstration more appropriately shows the productivity of the various diminutive meanings *relative to another*, i.e. diminutive formations in the categories of smallness, hypocoristics, pejoration, specialisation and ambiguous cases. For instance, the numbers in the column of 'Observed frequencies (Raw)' demonstrate the total sums of the lexemes' raw frequencies respective of each semantic category, i.e. the raw frequencies for all lexemes expressing e.g. smallness can be found in the 'Smalln.' column. The Spanish lexemes are

found under Corpus 1 (*CDE*) and the English lexemes under Corpus 2 (*COHA*)[37]. The corresponding relative frequencies have been calculated on the respective raw frequencies.

Table 4.10: Classification of English and Spanish diminutives according to semantic denotation

Classification of the Semantic Distribution of Diminutive Lexemes				
	Observed frequencies (Raw)		Normalised frequencies (PMW)	
Semantic Category	Corpus 1 (*CDE*)	Corpus 2 (*COHA*)	Corpus 1 (*CDE*)	Corpus 2 (*COHA*)
Smallness	5,432	33,966	1313.3	83.6
Hypocoristics	1,992	56,806	481.6	139.8
Pejoration	5,415	20,053	1309.2	49.4
Specialisation	19,193	44,800	4640.2	110.3
Ambiguous	5,969	50,547	1443.1	124.4
TOTAL	38,001	206,172	9187.4	507.5
Total size of corpus	41,362,172[38]	406,232,024[39]	-	-

Findings

As can be seen in table 4.10, the normalised frequencies (pmw) for each of the Spanish diminutive categories are markedly higher than the corresponding categories in English. This finding temporarily confirms the indication by previous research which characterises Spanish as a productive diminutive language and English as an unproductive one. The productivity or use of synthetic diminutives, through this semantic distributional perspective, is in total higher in Spanish than in English. The total relative frequencies (pmw) of all of the Spanish diminutive categories versus the corresponding English ones are markedly higher in Spanish, demonstrating an 18 times higher frequency in Spanish (9187.4 PMW) than in English (507.5 PMW). To be sure that the findings can be trusted, i.e. to discover whether the differences are statistically significant, a statistical significance test of the same values will be undertaken in

[37] The large differences that can be seen between the Corpus 1 and Corpus 2, in table 4.6, demonstrate the difference of size between the two corpora, and are not representative of the productivity of the various categories in the two languages.

[38] The total size of *CDE* was recently updated to include 100 times more words than it previously had, which is reflected in the current number. As such, it can be considered an open corpus as more data will most likely be added in the future.

[39] The total size of COHA has been the same since 2009, which is when it was created. Typically, historical corpora are closed as "there can be no further input to an area" (Natural Language Processing 2009. [online] http://language.worldofcomputing.net/linguistics/introduction/what-is-corpus.html)

the next part of the analysis. First, however, we must take a closer look at what these differences involve and demonstrate some examples from the corpora of the most productive diminutive categories for each language. The most productive diminutive category for each language, as demonstrated in table 4.10, will be discussed (in contrast to the other language), i.e. specialised diminutives for Spanish and hypocoristics for English.

The special productivity of Spanish specialised diminutives

If we compare the five semantic categories within Spanish in terms of relative frequency (table 4.10), we can see that the group of specialised diminutives is the most productive diminutive category. The diminutive formations pertaining to the group of specialised diminutives together demonstrate a relative frequency of 4640.2 pmw. As previously discussed, the primitive bases of specialised diminutives behave differently from the rest of the diminutive categories included in this analysis, i.e. they acquire a new, specialised meaning through the addition of a diminutive suffix, instead of denoting an emotive sense. Table 4.11 provides some examples of the specialised diminutives found in the *CDE*, illustrated together with the respective meaning changes that have occurred. The meanings of the primitives have been found in Gooch (1967: 90; 117; 152):

Table 4.11: Examples of Spanish specialised diminutives from *CDE* and their corresponding primitive meanings

Specialised diminutive from corpus	**Primitive base**
peseta(s) 'peseta(s)'	*peso* 'weight'; peso (coin)'
bolsillo(s) 'pocket'	*bolso* 'bag, handbag'
libreta(s) 'note-book'	*libro* 'book'
tortilla(s) 'omelette'	*torta* 'bun, cake, pastry'

Interestingly, we can see from the examples of specialised diminutives in 4.11 (these particular lexemes have deliberately been chosen to show this[40]) that the two most productive diminutive suffixes in the production of specialised diminutives are the suffixes *-illa* and *-eta*. The two feminine suffixes form more than half of the lexemes categorised as specialised

[40] See Appendix A/B for all formations of specialised diminutives, which will demonstrate the same that is claimed here.

diminutives, i.e. *-illa* with 35.7%[41] and *-eta* with 26.5% of the formations. The next suffixes in line can only be said to be relatively productive in comparison, i.e. *-ín* (12.2%), *-illo* (10.2%) and *-ete* (8.2%). Thus, we can see that feminine suffixes (*-illa* and *-eta*) are markedly more productive than masculine suffixes (*-ín*, *-illo* and *-ete*) in the formation of specialised diminutives. A few illustrative examples of this corpus finding will follow here (18-21):

(18) El aire se llena de alegría africana. Cada vehículo tiene su ruido o *campanilla* de anuncio, sonante y pueril como idioma de chiche.
[The air filled with African joy. Every car has its noise or *small bell* of notice, sounding and childlike as the language of baby food.]

(Fiction, 1900s, *CDE*)

(19) Creo que son dos las razones. Por un lado, el fortalecimiento de la *guerrilla* ha captado muchos de los cuadros jóvenes de estas organizaciones.
[I think there are two reasons for it. On the one hand, the strengthening of the *guerrilla warfare* has captured many of the young paintings of these organizations.]

(News, 1900s, *CDE*)

(20) El hombre manoteó al responder: - Claro, en el *bolsillo*; era mío. Estaba próximo a perder el control de sus nervios.
[The man gesticulated as he responded: - Of course, the *pocket*; was mine. He was close to losing his nerves.]

(Fiction, 1900s, *CDE*)

(21) Estas *mesetas* basálticas han sido creadas por volcanes.
[These basaltic *tables* were created by volcanoes.]

(Academic, 1900s, *CDE*)

The majority of the hits for each of the four specialised diminutives exemplified here, i.e. *campanilla(s)*, *guerrilla(s)*, *bolsillo(s)* and *meseta(s)* occur either in genres of Fiction, News or Academic texts. This observation confirms the role of specialised diminutives to denote a specialised meaning different from that of the primitive base, i.e. in a non-emotive way. Therefore, specialised diminutives will often be found in both informal and formal types of language, acting as non-emotive derivational derivatives, to apply the Spanish term. However, that said, the Fiction and News genres contain language of varying degree of formality, perhaps leaning towards more informal types of language (e.g. indirect speech, movie conversations, debates, etc.).

[41] 35.7% of the endings used to form specialised diminutives.

More in detail, the formations of *bolsillo(s)* are almost exclusively found in the Fiction genre of *CDE* (71.07 pmw). The example of *bolsillo* in (20) denotes 'pocket', whose meaning is also applied for *bolsillo* in Gooch (1967: 90). Only an insignificant amount of *bolsillo* formations are found in other the corpus genres, e.g. the Academic genre (1.80 pmw). Similarly, the majority of the formations of *campanilla* are found in the Fiction genre of *CDE*, as demonstrated by the example of *campanilla* 'small bell' in (18)[42]. The specialised diminutive *guerrilla(s)* is most frequently found in the News section of *CDE* (53.19 pmw.). The illustrated example of *guerrilla* in (19) refers to 'guerrilla warfare' occurring in the News section. Furthermore, another portion of *guerrilla*-formations is also found the Academic section of the corpus (30.80 pmw). These findings are not surprising, as one would logically assume to find the lexeme *guerrilla* in those two contexts, as wars typically receive much publicity from the media, and later are written about in history books. The primitive of *guerrilla* is *guerra* 'war'[43]. Furthermore, the instances of the specialised diminutive *meseta(s)* occur in the Academic section of the *CDE*. The example of *mesetas* in (21) is applied in the meaning of "tables" or "tablelands", which Gooch (1967) similarly describes as the most common function for this specialised diminutive (Gooch 1967: 152).

To sum up, based on the findings for the four lexemes looked at here, the overall trend for specialised diminutives seems to be that they tend to occur in the more "formal" or "serious" genres of the corpus, such as the Academic, News or Fiction genres. None of the four specialised diminutives looked at here occur in the genre of Spoken language. This trend appears to agree with the type of diminutive category they represent, i.e. in denoting a specialisation or change of meaning, instead of an emotive meaning.

The hypocoristics category

The most productive diminutive category of the English data is the hypocoristics category. The diminutives included in this category demonstrate a frequency of 139.8 pmw distributed on 27 lexemes. Hypocoristics typically occur in informal speech, filling the function of markers of affection and familiarity. In the data analysed, the two formations of *darling(s)* and *daddy(ies)* are drawn notice to as they demonstrate the highest productivity of the

[42] Noteworthy, the specialised diminutive *campanilla* has a "sister" in the diminutive lexeme of *campanita* 'e.g. (cute) little bell', which is not part of the analysis of this study, but this lexeme instead expressed the more typical diminutive sense of smallness or hypocorism (Gooch 1989: 111).

[43] A sister to *guerrilla* can be found in *Guerrita* 'young Guerra', used as a surname (Gooch 1967: 114).

lexemes in this diminutive category, the former, *darling(s)*, showing a relative frequency of 27.6 pmw and the latter, *daddy(ies)*, a relative frequency of 25.9 pmw. The two hypocoristics are demonstrated here with two examples (22-23) from the English corpus:

(22) "Would you like to sleep with me tonight, *darling*?" "But Mother, of course!"

(Fiction, 1940, *COHA*)

(23) Murderer, she whispers. I am too frail to flee. I see you watchin' your *daddy* and me. She pins me to the wall. You don' say nothin'.

(Fiction, 1991, *COHA*)

The two hypocoristic formations in (22) and (23) are both found in the Fiction genre of the English corpus, demonstrating the informality of speech that characterises the texts of this corpus genre. It should be noted that even though the English hypocoristics category is the most frequent of the English categories (139.8 pmw), it still demonstrates a much lower frequency that the Spanish hypocoristics category (481.6 pmw), which again is the least productive of all five Spanish diminutive categories. That says something about the difference of productivity in general between English and Spanish.

Interestingly, however, separate comparisons of individual hypocoristics in English and Spanish reveal that the English hypocoristics demonstrate higher frequencies than the Spanish ones, e.g. *darling* (27.6 pmw), *daddy* (25.9 pmw) vs. *casita* '(delightful) little house' (10.8 pmw), *Paquita* 'hypoc. for Francisca' (10.6 pmw). This finding is interesting because it reveals that despite the fact that the English language forms fewer diminutive lexemes than Spanish in general, the ones that are used in English, are by contrast to Spanish hypocoristics, very productive. It is, in fact, peculiar that a language so productive on diminutive suffixation must see itself beaten by a much less prominently synthetic diminutive language such as English.

Now that the various semantic categories have been compared to each other through their normalised frequencies, it is time to test if the observed differences of table 4.10 are statistically significant.

Statistical findings

The testing of statistical significance of the relative frequencies from table 4.10 will be carried out mainly through the method of log likelihood (LL) while the *p*-value test and Bayes factor

will be applied in addition (see 4.4.1). The differences in relative frequency demonstrated in table 4.10, showed large differences between English and Spanish in terms of the semantic distribution of diminutives. The method of LL will now be used to test the statistical significance of the differences indicated by the relative frequencies (see table 4.10). The relative frequencies are simply not sufficient to demonstrate whether the differences between the two datasets are significant or simply due to chance. As such, the log likelihood demonstration that will be carried out will help to answer the question posed in RQ2, namely:

> In terms of semantic distribution, which diminutives are more innovative – the English or the Spanish?

Table 4.12 presents the results of the statistical testing of the distribution of semantic categories in English and Spanish. The observed frequencies, log likelihood scores, the *p*-values and effect sizes for the five different semantic categories are demonstrated in the table. The LL scores range from 112.83 to 21931.11, which are all higher than 3.84, the value that has to be surpassed to pass as statistically significant. This means that the differences in semantic distribution between the two languages are significant, and cannot be ruled a coincidence, as explained by Levshina (2015). Spanish is convincingly more productive in diminutive use of all meanings in comparison to English. Similarly, the probability value *p* confirms this, as all the *p*-value scores are below 0.0001, indicating that it is 99.99% certain that the observed differences are not due to chance. To control these findings, the effect sizes have been included. The effect sizes show that there is a strong relationship between the different variables, i.e. that the variables differ greatly. Furthermore, the largest difference per category between English and Spanish is found in the category of specialised diminutives, which can boast of an LL score as high as 21931.11. The lowest LL score is 112.83 for the ambiguous category [44], signifying that this is the category that is most similar in use in English and Spanish.

[44] But is it really a category? Isn't that where you put the words that are difficult to categorize?

Table 4.12: Statistical significance scores for the semantic categories of diminutives in English and Spanish

Semantic Category	Observed Frequencies		Log likelihood (LL)	*P*-value	Bayes
	Corpus 1 (CDP)	Corpus 2 (COHA)			
Smallness	5,432	33,966	855.63	0.0001	835.71
Hypocoristics	1,992	56,806	3102.63	0.0001	3082.71
Pejoration	5,415	20,053	3326.00	0.0001	3306.08
Specialisation	19,193	44,800	21931.11	0.0001	21911.19
Ambiguous	5,969	50,547	112.83	0.0001	92.91

Thus, research question two "In terms of semantic distribution, which diminutives are more innovative – the English or the Spanish?" has been answered. It can be confirmed that the latter language as expected is the more innovative, and thus also the most productive of the two.

Overuse and underuse

Another factor which is indicated by the LL calculator, but which has not been included in table 4.12, relates to the fact that some semantic categories demonstrate overuse/underuse in C1 (*CDE*) relative to C2 (*COHA*). In four of five categories, namely in the categories of smallness, pejoration, specialisation and ambiguity, the Spanish corpus demonstrates overuse. This means that these four diminutive meanings are overused in the Spanish corpus in comparison to the English corpus.

Interestingly, however, the hypocoristics category demonstrates underuse in the Spanish corpus relative to the English corpus. This finding means that the relative frequency between the two languages is less than what we would expect when we compare the frequencies of the two corpora. Spanish diminutives in *CDE* are still more frequently expressing affection than the English diminutives in *COHA*, but one would have had reason to expect an even higher number of hypocoristics in the Spanish corpus (relative to what was going to be found in the English corpus), than was actually the case.

This finding is surprising because the category of hypocoristics represents one of the most basic diminutive types, and should occur more frequently in a diminutive-rich language such as Spanish. English diminutives, on the other hand, are expected to be frequent in the production of hypocoristics, as unproductive diminutive languages typically are best represented by the basic diminutive meanings. Some explanations for this finding may be due

to a difference in the composition of the two corpora's genres. Perhaps the genres of the English corpus contain more informal language than the genres of the Spanish corpus, despite the fact that the *CDE* contain a Spoken genre which *COHA* does not. A comparison of the various genres in the two corpora demonstrates the following:

Table 4.13: Composition of the two corpora applied in the study

COHA	CDP
Fiction (novels, movie and play scripts)	Fiction/Literature (novels, short stories)
Popular Magazines	Oral (interviews, transcripts)
Newspapers	News (articles)
Non-Fiction	Academic (articles)

The *COHA* Fiction section also includes movies in its repertoire, which the Spanish does not. Movies tend to apply much informal language, perhaps more so than novels, so that may be a reason for the rather odd finding. In order to check if this is really is the case, that the Fiction genre in *COHA* is the reason for the high number of hypocoristics turning up in the English data, one would have to study each hypocoristic in detail. This is not possible in the current study. A random sample of 500 for each of the 27 hypocoristic diminutives has instead been collected, which showed that the majority of the hypocoristics was from the *COHA* Fiction section. An inspection of each sentence to discover whether the sentence was taken from a novel or a movie was, however, not possible due to the limited scope of the study.

4.4.5 Semantic trends

This semantic distributional analysis of the use of English and Spanish synthetic diminutives in the last two hundred years has shown that Spanish is the more productive and innovative language of the two. Each of five diminutive categories (smallness, hypocoristics, pejoration, specialisation and ambiguous cases) considered in the semantic analysis demonstrated markedly higher relative frequencies for Spanish than the corresponding categories in English (see table 4.10). In addition, the statistical testing of the marked differences proposed by the relative frequencies established that the differences were extremely significant and cannot be ruled a coincidence (see table 4.12). This finding thus confirms indications by previous research, which this study aimed to test, characterising Spanish as a productive diminutive language and English an unproductive one. The hypothesis proposed at the start of the study has thus been confirmed.

The largest difference in semantic distribution between the two languages is seen in the category of specialised diminutives. The results from the relative frequency comparison and the statistical testing shows that Spanish is significantly more productive in the use and formation of this diminutive type (e.g. LL of 21931.11). Furthermore, in Spanish we can see that the two feminine diminutive suffixes *-illa* and *-eta* have been the two most productive suffixes within this diminutive category the last two hundred years. Also, there seems to be a trend for specialised diminutives to occur in the more "formal" or "serious" genres of the *CDE* such as the Academic or News genres. This trend appears to agree with the type of diminutive category they represent, i.e. in denoting a specialisation or change of meaning, instead of an emotive meaning.

The most productive diminutive category of the English data is the hypocoristics category (though beaten here, too, by Spanish). Interestingly, however, separate comparisons of hypocoristic lexemes in English and Spanish reveal that the English hypocoristics demonstrate higher frequencies than the Spanish ones, e.g. *darling* (27.6 pmw), *daddy* (25.9 pmw) vs. *casita* '(delightful) little house' (10.8 pmw), *Paquita* 'hypoc. for Francisca' (10.6 pmw). This finding is interesting because it reveals that despite the fact that the English language forms fewer diminutive lexemes than Spanish in general, the ones that are used in English are, in contrast to many Spanish hypocoristics, very productive. It is, in fact, peculiar that a language so productive on diminutive suffixation must see itself beaten in lexeme-lexeme comparisons by a much less prominently synthetic diminutive language such as English.

5 Conclusive remarks

5.1 Strengths and findings of the study

The aim of this thesis has been to provide an overview of the diminutive category in the English and Spanish languages, in terms of productive lexeme formations and the distribution of semantic denotations expressed across these two contrasting languages during the last two hundred years (c. 1800-2000). The study has aimed to shed light on the contrasts between the two different languages, i.e. the emotive language of Spanish and the less emotive English language.

The results presented in chapter 4 show the strength and usefulness of having applied a quantitative, and in particular, a corpus-based approach to the study of diminutives. This approach has enabled the study of large quantities of data and of diminutives in real life situations, i.e. in naturally occurring language, making it possible to make claims about the manifestation of diminutives in two contrasting languages. It has been possible to challenge the common myth that English is an unproductive language with respect to diminutives, a myth that stems from the firmly rooted tradition of introspection-based diminutive research. This study has, in fact, empirically shown that English is quite productive in the use of a few diminutive forms, i.e. the most frequent English diminutive lexemes, which, in fact, beat some of the Spanish diminutives in terms of relative frequency. This is particularly evident in the hypocoristics category, where the most productive English hypocoristics are more productive than several of the most productive Spanish hypocoristics (see table 4.3). That said, however, there is no doubt that Spanish is the more productive of the two, as this study has empirically proven. The hypothesis of this study has been confirmed through two diminutive investigations, which firstly, declared Spanish as more productive on diminutive formation, and secondly, the more innovative in the distribution of diminutive meanings. The overall conclusion of this study is thus that Spanish is vastly more productive in the formation of diminutive words, while English can be said to be productive in the use of the lexemes that exist in the English diminutive repertoire.

A second factor worth mentioning is the contrastive starting point of the discussion, i.e. contrasting English to Spanish – a well-developed and innovative diminutive language – which turned out to contribute to the finding of a new diminutive use in English, namely specialised diminutives. This diminutive type is well-established in Spanish diminution, but has previously not been acknowledged by English diminutive research. It is argued in this

study that specialised diminutives should be established as a diminutive category in addition to the already existing repertoire of diminutive meanings in English, i.e. the senses of smallness, hypocorism and pejoration. This study, furthermore, suggests that if we include specialised diminutives in the diminutives class, the productivity of the English diminutive category would be considered to be higher than is generally perceived today.

However, this study does not propose that the category of specialised diminutives will have an important role in the same way as it currently enjoys in Spanish. This study simply claims that such diminutives exist in English. It goes against nature for English to use derivation as the main word-forming device, as it prefers compounding to produce new words. In Spanish, derivation is the main word-forming device. Consequently, specialised diminutives are much more naturally formed within the language. This can be illustrated by the example of the primitive *mesa* 'table' which has two derivational derivatives, i.e. the pure diminutive *mesita* '(cute) little table' and the specialised diminutive *meseta* 'tableland'. The meaning of *meseta* is not emotive, and as such the diminutive function as any other word formed on a non-emotive suffix in Spanish. English has instead of applying derivation, formed a new compound word from *table* > *tableland*.

5.2 Limitations and suggestions for further research

Some limitations of this study are that the present study falls too short in the coverage of ambiguous meanings expressed by diminutives, the corpus data demonstrating many cases of ambiguity, which the quantitative approach applied, was not able to investigate. Future corpus-based studies should aim at applying qualitative approaches to the study of diminutives, as this would ensure deeper and more thorough investigations of how diminutives are used in language. We cannot be sure that a diminutive typically described to be used in a pejorative manner, as e.g. by Gooch (1967) or the *OED* (2016), is actually used in this way in all instances of the given lexeme in the corpus. Thus, the present study only indicates various typical meanings expressed by diminutives. Particularly in the case of ambiguous diminutives, which constitute a rather large portion of the diminutives of the present research, the limited scope of this study has not been able to guarantee a correct portrayal of use of the given diminutives. My study can at best only be tentative in the depiction of diminutive meanings expressed, as the focus has been on documenting the frequency and productive uses of diminutives.

Furthermore, if tagging becomes better in the future, it would allow the search of diminutives without manually having to sort the lexemes that express an emotion from common non-emotive words which end in the same letters as a diminutive suffix. Tagging should include semantic captions as well as the grammatical ones existing today.

Also a more precise study of diminutive meaning between the English and Spanish language pair should be aimed for in future research. Gooch's (1967) description of the various shades and nuances expressed by Spanish diminutives can most likely not be said to apply to English diminutives, since English is a less developed diminutive language. In order to investigate the semantic distribution of diminutives between the current language pair, fewer lexemes must be selected, and a qualitative approach employed.

It is my hope that future contrastive diminutive studies will be able to advance the present study's discussion of diminutives, importantly from a corpus perspective, applying a more qualitative-based approach to diminutives. This would enable a more in-depth study of diminutive meaning, and would improve the quality of the study of the ambiguous nature of diminutives.

As a leaving note, the present study has first and foremost aimed to be of theoretical use, i.e. to contribute to the lack of empirical research existing in the field of contrastive word formation. It is my hope that this has been accomplished by this study. It was claimed in the introduction to this thesis that contrastive word formation lies at the crossroads between contrastive linguistics, lexicology and morphology. As such, the present study may provide insights into the contrastive use of the morphological process of diminution which can be applied for the learning and teaching of English and Spanish as a foreign language.

Bibliography

Andor, J. (2005). 'A lexical semantic-pragmatic analysis of the meaning potensials of amplifying prefixes in English and Hungarian. A corpus-based case study of near-synonomy', in Proceedings from the Corpus Linguistics Conference Series 1 (1).

Arnaud, P.J.L and Renner, V. (2008). 'English and French $[NN]_N$ compounds. A categorial, morphological and semantic comparison'. Paper presented at the *5th International Contrastive Linguistics Conference*, 7-9 July, Leuven, Belgium.

Ascoop, K. and Leuschner, T. (2006). "Affixoidhungrig? Skitbra! Comparing affixoids in German and Swedish", *Sprachtypologie und Universalienforschung* 59 (3), 241-252.

Augustyn, R. and Gniecka, A. (2011). 'Irony behind diminutives: a cognitive linguistic analysis of popular technical terms', in *Lublin Studies in Modern Languages and Literature* 35, 31-49.

Bauer, L. (1978). *The grammar of nominal compounding with special reference to Danish, English and French*. (Odense University Studies in Linguistics 4.) Odense: Odense University Press.

Bauer, L. (1983). *English Word-Formation*. Cambridge: University Press.

Bauer, L. (2001). *Morphological productivity*. Cambridge studies in linguistics 95. Cambridge: Cambridge University Press

Beard, R. and B. Szymanek. (1988). *Bibliography of morphology 1960–1985*. Amsterdam: Benjamins.

Bertrand, C. (1986). 'Quelques remarques sur les adverbes francais en *-ment* et leur rapport avec les adverbes anglais en *-ly*', *Meta* 31 (2), 179-203.

Booij, G. & R. Lieber. (2004). "On the paradigmatic nature of affixal semantics in English and Deutsch", *Linguistics*, 42(2), 327-357.

Boyd-Bowman, P. (1955). 'Cómo obra la fonética infantil en la formación de los hipocorísticos', in *Nueva Revista de Filología Hispánica* 9, 337-366.

Bradley, H. (1908). *The Making of English*. London: Macmillan & Co. Ltd.

Bulloch, J. M. (1921/1970). 'The delight of the Doric in the diminutive', in W. A. Craigie (ed.). *The Scottish Tongue* College Park, Md.: McGrath.

Cannon, G. (1987). *Historical Change and English Word-Formation*. New York: Peter Lang.

Charleston, B. (1960). *Studies on the Emotional and Affective Means of Expression in Modern English*. Bern: Francke Verlag Bern.

Chesterman, A. (1998). *Contrastive functional analysis*. Amsterdam: Benjamins.

Collins Spanish Dictionary (2005). London: HarperCollins Publishers.

Cvilikaitė, J. (2007). "The cases of non-equivalence between English and Lithuanian: A corpus-based solution". *Kalbotyra* 57 (3), 46-55.

Davies, M. (2002). *Corpus del Español: 100 million words, 1200s-1900s*. Available online at http://www.corpusdelespanol.org

Davis, M. (2002). *Corpus of Historical American English, 1810-2009*. Available at http://corpus.byu.edu/coha/

Davis, M. (2010). 'Creating Useful Historical Corpora: A Comparison of CORDE, the Corpus del Español, and the Corpus do Português', in Enrique-Arias, A. (ed.). *Diacronía de las lenguas iberorromances: nuevas perspectivas desde la lingüística de corpus*. Frankfurt/Madrid: Vervuert/Iberoamericana, 137-66.

Dossena, M. (2012). 'Vocative and diminutive forms in Robert Louis Stevenson's fiction: A corpus-based study, in *International Journal of English Studies* 12 (2), 1-17.

Dressler, W. U. and Barbaresi, M. L. (1994). *Morphopragmatcs: Diminutives and Intensifyers in Italian, German and other Languages*. Berlin: Mouton de Gruyter.

Fedorowicz-Bacz, B. (1977). 'Syntactic function of negative prefixes in English and Polish participial adjectives', in *Papers and Studies in Contrastive Linguistics* 6, 111-122.

Fishman, J. (1964). 'Language Maintenance and Language Shift as a Field of Inquiry', in *Linguistics* 9, 32-70.

Funk, W.-P. (1986). 'Towards a definition of semantic constraints on negative prefixation in English and German', in Kastovsky, D. and A. Szedek (eds.), *Linguistics across*

historical and geographical boundaries. Volume 2: Descriptive, contrastive and applied linguistics. Berlin: Mouton de Gruyter, 877-889.

Gast, V. and Hüning, M. (2008). 'V+N-compounds in West-Germanic'. Paper presented at the *International Contrastive Linguistics Conference 5*, 7–9 July, Leuven, Belgium.

Goddard, C. and A. Wierzbicka (2008). 'Universal human concepts as a basis for contrastive linguistic semantics', in de los Angeles Gomez Gonzalez, M., J. L. Mackenzie and E. M. Gonzalez Alvarez (eds.). *Current trends in contrastive linguistics. Functional and cognitive perspectives*. Amsterdam: Benjamins, 205-226.

Gooch, G. (1967). *Diminutive, augmentative and pejorative suffixes in modern Spanish: (a guide to their use and meaning).* Oxford: Pergamon Press.

Granger, S. and Rayson, P. (1998). 'Automatic profiling of learner texts', in Granger, S. (ed.). *Learner English on Computer.* London: Longman, 119-131.

Guevera, E. and Scalise, S. (2004). *V-compounding in Dutch and Italian*. (Cuadernos de Linguistica del Instituto Universiatrio Ortega y Gasset, Madrid.)

Haas, M. (1972). 'The Expression of the Diminutive', in Smith, M. E. (ed.). *Studies in Linguistics: In honor of George L. Trager*. The Hague: Mouton, 148-152.

Hammer, J. H. and F.A. Rice. (1965). *A bibliography of contrastive linguistics*. Washington, D. C.: Center for Applied Linguistics.

Heyvaert, L. (1998). 'Non-agentive deverbal *-er* nominalization in English and Dutch: A contrastive analysis', in *Languages in Contrast* 1 (2), 211-243.

Huddleston, R. and Pullum, G. K. (2002). *The Cambridge Grammar of the English Language.* Cambridge: Cambridge University Press, 1565-1619 and 1621-1721.

James, C. (1980). *Contrastive analysis*. London: Longman.

Jespersen, O. (1942*). A modern English grammar on historical principles*. London: George Allen.

Jurafski, D. (1996). 'Universal Tendencies in the Semantics of the Diminutive', in *Journal of the Linguistic Society of America*, 72 (3), 533-78.

Kačmárová, A. (2010). ‘A short survey of diminutives in Slovak and English’, in Kačmárová, A. (ed.), *English matters: a collection of papers by the Department of English language and literature faculty*. Prešov: Prešovská univerzita v Prešove, 17-21.

Katamba, F. (2005). *English Words: Structure, History, Usage*. Oxford: Routledge.

Kilgarriff, A. (1996). *Why chi-square doesn't work; and an improved LOB-Brown comparison.* ALLCACH Conference, June 1996, Bergen, Norway.

Kin Quah, C. (1999). ‘Issues in the translation of English suffixes into Malay’, in *Meta* 44 (4), 604-616.

Kryk-Kastovsky, B. (2000). ‘Diminutives: an interface of word-formation, semantics and pragmatics’, in Dalton-Puffer, C. and N. Ritt (eds.). *Words: Structure, Meaning, Function. A Festschrift for Dieter Kastovsky*. Berlin: Mouton de Gruyter.

Krzeszowski, T. P. (1990). *Contrasting languages. The scope of contrastive linguistics*. Berlin: Mouton de Gruyter.

Lakoff, G. (1987). *Women, Fire and Dangerous Things: What Categories Reveal about the Mind*. Chicago: Chicago University Press.

Lang, M. F. (1990). *Spanish Word Formation*. London: Routledge.

Lefer, M.-A. and Cartoni, B. (2011). ‘Prefixes in contrast: Towards a meaning-based contrastive methodology for lexical morphology’, in *Languages in Contrast* 11 (1), 86-104.

Lefer, M.-A. (2011). ‘Contrastive Word-Formation Today: Retrospect and Prospect’, in *Poznań Studies in Contemporary Linguistics* 47 (4), 645-682.

Leisi, E. (1969). *Das heutige Englisch. Wesenszüge und Probleme.* Heidelberg: Winter.

Levshina, N. (2015). *How to do Linguistics with R: Data exploration and statistical analysis*. Amsterdam: John Benjamins.

Lipski, J. M. (1990). *The language of the islenos: vestigial Spanish in Louisiana*. Baton Rouge: Louisiana State University Press.

Marchand, H. (1969). *The Categories and Types of Present-Day English Word-Formation*. München: Beck.

McEnery, T, and Hardie, A. (2012). Corpus Linguistics. Cambridge: Camrbridge Textbooks in Linguistics.

Náñez Fernández, E. (1973). *El diminutivo. Historia y funciones en el español clásico y moderno*. Madrid: Editorial Gredos, S. A.

Natural Language Processing (2009). What is corpus? Accessed online at http://language.worldofcomputing.net/linguistics/introduction/what-is-corpus.html) [Accessed April 4, 2016].

Narváez, R. A. (1970). *An Outline of Spanish Morphology*. Saint Paul, Minnesota: EMC Corporation.

Nieuwenhius, P. (1985). *Diminutives*. Unpublised PhD thesis. Edinburgh: University of Edinburgh.

Oakes, M. (1998). *Statistics for Corpus Linguistics*. Edinburgh: Edinburgh University Press.

Oxford English Dictionary (2016). Oxford: Oxford University Press.

Paillard, M. (2000). *Lexicologie contrastive anglais-fracais. Formation des mots et construction du sens*. Paris: Ophrys.

Plag, I., C. Dalton-Puffer and H. Baayen (1999). 'Morphological productivity across speech and writing', in *English Language and Linguistics* 3 (2), 209-228.

Poynton, C. (1989). 'Terms of address in Australian English', in Collins P. and D. Blair (eds.). *Australian English. The Language of a New Society*. St. Lucia, Queensland: University of Queensland Press, 55-69.

Prieto, P. (1992). 'Truncation processes in Spanish', in *Studies in the Linguistic Sciences* 22, 143-158.

Quirk, R., S. Greenbaum, G. Leech and J. Svartvik (1972). *A Grammar of Contemporary English*. London: Longman.

Quirk, R., S. Greenbaum, G. Leech and J. Svartvik (1985). *A Comprehensive grammar of the English language*. London: Longman.

Rabadán, R., B. Labrador and N. Ramón. (2005). 'Putting meaning into words: English '-ly' adverbs in Spanish translations'. Paper presented at the *International Contrastive Linguistics Conference 4*, 19-23 Sept. 2005, Compostela, Spain.

Rayson. P., G. Leech and H. Hodges (1997). 'Social differentiation in the use of English vocabulary: some analyses of the conversational component of the British National Corpus', in *International Journal of Corpus Linguistics* 2 (1), 133-152.

Rayson P. and Garside R. (2000). 'Comparing corpora using frequency profiling', in *CompareCorpora '00 Proceedings of the Workshop on Comparing Corpora*. Stroudsburg, PA: Association for Computational Linguistics, 1-6.

Real Academia Española y Asociación de Academias de la Lengua Española. Versión beta Nueva gramática de la lengua española (2009). [Edición en línea (www.rae.es). Acceso: 2016-04-20].

Real Academia Española y Asociación de Academias de la Lengua Española (2010). *Nueva gramática de la lengua española. Morfología y sintaxis*. Madrid: Espasa. http://www.rae.es/obras-academicas/gramatica/nueva-gramatica/nueva-gramatica-morfologia-y-sintaxis#sthash.XxP9MSym.dpuf

Real Academia Española (2014). *Diccionario de la lengua española*. Madrid: Espasa [online] http://www.rae.es/obras-academicas/diccionarios/diccionario-de-la-lengua-espanola#sthash.lBSyNoZU.dpuf

Rosiak, K. (2013). 'From Synthetic to Analytic? The Changing Use of Diminutive Expressions in Welsh.' In: Ahlqvist, Anders; O'Neill, Pamela (eds.) *Celts and their Cultures at home and abroad. A Festschrift for Malcolm Broun.* Sydney: The Celtic Studies Foundation, the University of Sydney, 289-309.

Rotzoll, E. (1910). *Die Deminutivbildungen im Neuenglischen unter besonderer Berücksichtigung der Dialekte*. Heidelberg: Winter.

Rusek, M. (2005). 'That English has no diminutives is a common myth - based on Klaus P. Schneider's book "Diminutives in English"', Munich, GRIN Verlag, http://www.grin.com/en/e-book/77184/that-english-has-no-diminutives-is-a-common-myth-based-on-klaus-p-schneider-s

Sajavaara, K. and J. Lehtonen. (1975). *A select bibliography of contrastive analysis.* (Jyväskylä Contrastive Studies, 1. Reports from the Department of English, University of Jyväskylä, No. 1.)

Savickiene, I. and Dressler, W. U. (eds.)(2007). *The Acquisition of Diminutives: A cross-linguistic perspective*. Amsterdam: John Benjamins.

Scalise, S. (2005). *Generative Morphology*. Riverton, USA: Foris Publication.

Schneider, K. P. (2003). *Diminutives in English*. Tübingen: Max Niemeyer .

Sifianou, M. (1992). 'The use of diminutives in expressing politeness: Modern Greek versus English', in *Journal of Pragmatics* 17 (2), 155-173.

Sinclair, J. (1987). *Collins COBUILD English Language Dictionary*. London, Glasgow, Stuttgart: Collins/Klett.

Stein, G. (1973). *English word-formation over two centuries. In honour of Hans Marchand on the occasion of his sixty-fifth birthday*. Tübingen: Narr.

Steriopolo, O. (2015). Syntactic variation in expressive size suffixes. A comparison of Russian, German and Spanish, in *SKASE Journal of Theoretical Linguistics*, 12 (1), 2-20.

Strang, B. (1970). *A History of English*. London: Methuen.

Stubbs, M. (2002). *Words and Phrases. Corpus studies of lexical semantics*. Oxford: Blackwell.

Sussex, R. (2004). '*Abstand, Ausbau*, Creativity and Ludicity in Australian English', in *Australian Journal of Linguistics* (24), 3-19.

Szymanek, B. (1989). *Introduction to morphological analysis*. Warszawa: PWN.

Taylor, J. R. (2003). *Linguistic Categorization* Oxford Textbooks in Linguistics. Oxford: OUP Oxford.

Terrell, Tracy David and Salgués de Cargill, M. (1979). *Linguistica Aplicada: A la enseananza del Espanol a Anglohablantes*. New York: John Wiley and Sons.

Togini-Bonelli, E. (2001). *Corpus Linguistics at Work*. Amsterdam: Benjamins

Whitley, M. S. (1986). *Spanish/English Contrasts: A Course in Spanish Linguistics*. Washington D. C.: Georgetown University Press.

Williams, I. (2006). 'Towards a target-oriented model for quantitative contrastive analysis in translation studies. An exploratory study of theme-rheme structure in Spanish-English biomedical research articles', in *Languages in Contrast* 6 (1), 1-45.

Wierzbicka, A. (1984). Diminutives and Depreciatives. Semantic representation for Derivational Categories, in *Quaderni di Semantica* 9, 123-130.

Wierzbicka, A. (1985). ‘Different cultures, different languages, different speech acts’, in *Journal of Pragmatics* 9, 145-178.

Wierzbicka, A. (1991). Cross-Cultural Pragmatics. Berlin/New York: Mouton de Gruyter.

Appendices

<u>Appendix A</u>: Semantic classification of English and Spanish diminutives Findings from *CDE* and *COHA*.

Corpus data from *CDE*:

Diminutive type	Spanish lexemes (CDE)	Raw freq.
Diminutive (Smallness) -small in size -(slightly) affectionate -favourable implication -(augmentative)-pejorative implications) -(ironically)	*señorito|señorita(s)* (smalln./pej./spec.)	2097
	chiquillo/a(s) (smalln./pej.)	801
	casita(s)[45] (smalln./hypoc./spec.)	447
	festín(es) (spec./smalln./pej.)	338
	botín(es) (spec./smalln./pej.)	305
	mesita(s) (smalln./hypoc.)	289
	pajarito(s) (smalln./hypoc./pej.)	206
	corrillo(s) (smalln./pej.)	160
	camarín(es) (spec./smalln./pej.)	159
	hijita(s) (smalln./hypoc.)	144
	pajarillo(s) (smalln./often pej.)[46]	133
	calcetine(s) (spec./smalln./pej.)	119
	ratito(s) (smalln./hypoc./pej.)	119
	maletín(es) (spec./smalln./pej.)	115
	Total freq. of all dims. smalln.	**5432**
Hypocoristics	*casita(s)* (smalln./hypoc./spec.)	447
	Paquita	438
	mesita(s) (smalln./hypoc.)	289

[45] *Casita* can also be used ironically, either for humoristic effect, or negatively, to cause insult. When it is used negatively, it does naturally express some kind of pejorative sense. However, pragmatic senses, such as irony or sarcasm, are beyond the scope of this rather quantitative study. The same goes for *mesita*.

[46] Cf. *pajarito* '(sweet) little bird'

	Frasquito(s)	220
	pajarito(s) (smalln./hypoc./pej.)	206
	hijita(s) (smalln./hypoc.)	144
	Enriqueta	129
	ratito(s) (smalln./hypoc./pej.)	119
	Total freq. of all dims. hypoc.	**1992**
Specialisation -Some also often have diminutive and also pejorative implications (blue colour*)	*peseta(s)*	3234
	señorito|señorita(s) (smalln./pej./spec.)	2097
	bolsillo(s)	1604
	pasillo(s)	786
	cigarrillo(s)	743
	*juguete(s)** (spec./pej.)	743
	guerrilla(s)	588
	meseta(s)	462
	banquete(s)	455
	*casita(s)** (smalln./hypoc./spec.)	447
	carreta(s)	439
	campanilla(s)[47]	412
	trompeta(s)	363
	cuadrilla(s)	348
	careta(s)	343
	*festín(es)** (spec./smalln./pej.)	338
	ventanilla(s)	325
	botín(es) (spec./smalln./pej.)	305
	puntilla(s)	278
	camiseta(s)	227
	madrina(s)	226

[47] Cf. *campanita* '(cute) little bell' (Goosh 1967: 111)

	pastilla(s)	220
	boletín(es)	220
	camioneta(s)	215
	neblina(s)	194
	ramillete(s)	184
	maceta(s)	174
	libreta(s)	172
	camarín(es) (spec./smalln./pej.)	159
	tortilla(s)	157
	banqueta(s)	153
	banquillo(s)	152
	pandilla(s)	151
	cuartilla(s)	151
	muleta(s)	149
	cepillo(s)	149
	corneta(s)	145
	papeleta(s)	144
	cabecilla(s)	144
	patilla(s)	141
	taquilla(s)	140
	buhardilla(s	139
	barbilla(s)[48]	137
	barandilla(s)	132
	boquete(s)	121
	calcetine(s) (spec./smalln./pej.)	119
	bombilla(s)[49]	118

[48] Cf. *Barbita* ‘small beard’

	mantequilla(s)	116
	*maletín(es)** (spec./smalln./pej.)	115
	Total freq. of all spec.	**19074**
Pejorative	*señorito|señorita(s)* (smalln./pej./spec.)	2097
	chiquillo/a(s) (smalln./pej.)	801
	juguete(s) (spec./pej.)	743
	festín(es) (spec./smalln./pej.)	338
	botín(es) (spec./smalln./pej.)	305
	pajarito(s) (smalln./hypoc./pej.)	206
	corrillo(s) (smalln./pej.)	160
	camarín(es) (spec./smalln./pej.)	159
	pajarillo(s) (smalln./often pej.)[50]	133
	marica(s)	120
	calcetine(s) (spec./smalln./pej.)	119
	ratito(s) (smalln./hypoc./pej.)	119
	maletín(es) (spec./smalln./pej.)	115
	Total freq. of all dims. pej.	**5415**
Ambiguous	*señorito|señorita(s)* (smalln./pej./spec.)	2097
	chiquillo/a(s) (smalln./pej.)	801
	juguete(s) (spec./pej.)	743
	casita(s) (smalln./hypoc./spec.)	447
	festín(es) (spec./smalln./pej.)	338
	botín(es) (spec./smalln./pej.)	305
	mesita(s) (smalln./hypoc.)	289
	corrillo(s) (smalln./pej.)	160

[49] Cf. *bomita* ‘small pump or bomb’

[50] Cf. Pajarito ‘(sweet) little bird’

	camarín(es) (spec./smalln./pej.)	159
	hijita(s) (smalln./hypoc.)	144
	pajarillo(s) (smalln./often pej.)[51]	133
	calcetine(s) (spec./smalln./pej.)	119
	ratito(s) (smalln./hypoc./pej.)	119
	maletín(es) (spec./smalln./pej.)	115
	Total freq. of all amb.	**5969**
Total size of corpus (1800s-1900s)	(41,362,172)	**41362172**

Corpus data from *COHA*:

Domain	**English lexemes (*COHA*)**	**Raw freq.**
Smallness	*cigarette(s)* (smalln./spec.)	17968
	pony(ies) (smalln./spec.)	6226
	pamphlet(s)	3859
	cookie(s) (smalln./spec.)??	3055
	buggy(ies) (spec./smalln.)	2674
	tablet(s) (spec./smalln.)	2559
	puppy(ies) (smalln./hypoc./pej.)	2339
	hamlet(s)	2026
	Dolly(ies) (smalln./hypoc./pej./spec)	1931
	shanty(ies)|*chanty*|*shantie*|*shantee* (smalln./pej./spec.)	1392
	*brownie(s)**[52] (spec./pej./smalln.)	1166

[51] Cf. *pajarito* '(sweet) little bird'

[52] Pejoration/hypocoristic(affection)/smallness

	sapling(s) (smalln./pej./spec.)	1117
	rivulet(s)	1085
	leaflet(s)	1029
	seedling(s)	939
	kindling(s) (smalln./spec.)	801
	booklet(s)	744
	starling(s) (spec./smalln.)	597
	islet(s)	546
	yearling(s) (smalln./spec.)	493
	pixie(s) (smalln./hypoc./spec)	480
	droplet(s)	433
	laddie(s) (hypoc./smalln.)	426
	stripling(s) (smalln./pej.)	409
	foundling(s) (smalln./pej./spec.)	376
	streamlet(s)	354
	circlet(s)	304
	Total freq. of all dims. smalln.	**33966**
Hypocoristics	*darling(s)*	11217
	daddy(ies)	10525
	mammy(ies)|mommy/ies)|mummy/mamy(ies)	2908
	buddy(ies) (hypoc./spec.)	2813
	bunny(ies)	2789
	granny(ies)	2438
	puppy(ies) (smalln./hypoc./pej.)	2339
	auntie(s)	2255
	missy(es) (hypoc./pej.)	2025
	hippy|hippie(s) (hypoc./pej.)	2000
	bully(ies) (hypoc./pej.)	1998
	Dolly(ies) (smalln./hypoc./pej./spec)	1931
	jockey(s)	1717

	shanty(ies)|*chanty*|*shantie*|*shantee* (smalln./pej./spec.)	1392
	pussy(ies) (hypoc./pej.)	1192
	sweetie(s) (hypoc./spec.)	819
	nanny(ies) (hypoc./pej.)	797
	dearie(s)	683
	panty(ies) (spec./hypoc./pej.)	652
	Charley(s)|*Charlie(s)* (spec./hypoc./pej.)	541
	Collie(s)|*collie(s)* (spec./hypoc.)	533
	lassie(s)|*Lassie* (hypoc./spec)	510
	pixie(s) (smalln./hypoc./spec)	480
	Archie(s) (hypoc./spec.)	448
	laddie(s) (hypoc./smalln.)	426
	Jimmy(ies)	385
	homey(s)|*homie(s)* (hypoc./spec.)	306
	Total freq. of all dims. hypoc.	**56806**
Pejoration	*missy(es)* (hypoc./pej.)	2025
	puppy(ies) (smalln./hypoc.?/pej.)	2339
	hippy|*hippie(s)* (hypoc./pej.)	2000
	bully(ies) (hypoc./pej.)	1998
	Dolly(ies) (smalln./hypoc./pej./spec)	1931
	shanty(ies)|*chanty*|*shantie*|*shantee* (smalln./pej./spec.)	1392
	pussy(ies) (hypoc./pej.)	1192
	brownie(s) (spec./pej./smalln.)	1166
	sapling(s) (smalln./pej./spec.)	1117
	nanny(ies) (hypoc./pej.)	797
	panty(ies) (spec./hypoc./pej.)	652
	dumpling(s) (pej./spec.)	545
	Charley(s)|*Charlie(s)* (spec./hypoc./pej.)	541

	weakling(s)	533
	stripling(s) (smalln./pej.)	409
	foundling(s) (smalln./pej./spec.)	376
	cockney(s)|*Cockney(s)* (spec./pej.)	345
	changeling(s) (pej./spec.)	311
	underling(s)	384
	Total freq. of all dims. pej.	**20053**
Specialisation	*cigarette(s)* (smalln./spec.)	17968
	pony(ies) (smalln./spec.)	6226
	cookie(s) (smalln./spec.)	3055
	buddy(ies) (hypoc./spec.)	2813
	buggy(ies) (spec./smalln.)	2674
	tablet(s) (spec./smalln.)	2559
	Dolly(ies) (smalln./hypoc./pej./spec)	1931
	trolley(ies)|*trolly*	1681
	shanty(ies)|*chanty*|*shantie*|*shantee* (smalln./pej./spec.)	1392
	*brownie(s)**[53] (spec./pej./smalln.)	1166
	sapling(s) (smalln./pej./spec.)	1117
	sweetie(s) (hypoc./spec.)	819
	kindling(s) (smalln./spec.)	801
	caddie(s)	687
	panty(ies) (spec./hypoc./pej.)	652
	starling(s) (spec./smalln.)	597
	dumpling(s) (pej./spec.)	545
	Charley(s)|*Charlie(s)* (spec./hypoc./pej.)	541
	Collie(s)|*collie(s)* (spec./hypoc.)	533

[53] Pejoration/hypocoristic(affection)/smallness

	lassie(s)|*Lassie* (hypoc./spec)	510
	yearling(s) (smalln./spec.)	493
	pixie(s) (smalln./hypoc./spec)	480
	Archie(s) (hypoc./spec.)	448
	foundling(s) (smalln./pej./spec.)	376
	cockney(s)|*Cockney(s)* (spec./pej.)	345
	changeling(s) (pej./spec.)	311
	homey(s)|*homie(s)* (hypoc./spec.)	306
	Total freq. of all dims. spec.	**44800**
Ambiguous	*cigarette(s)* (smalln./spec.)	17968
	pony(ies) (smalln./spec.)	6226
	cookie(s) (smalln./spec.)	3055
	buddy(ies) (hypoc./spec.)	2813
	buggy(ies) (spec./smalln.)	2674
	tablet(s) (spec./smalln.)	2559
	puppy(ies) (smalln./hypoc.?/pej.)	2339
	missy(es) (hypoc./pej.)	2025
	bully(ies) (hypoc./pej.)	1998
	Dolly(ies) (smalln./hypoc./pej.)	1931
	shanty(ies)|*chanty*|*shantie*|*shantee* (smalln./pej./spec.)	1392
	pussy(ies) (hypoc./pej.)	1192
	*brownie(s)**[54] (spec./pej./smalln.)	1166
	*sapling** (smalln./pej.)	1117
	kindling(s) (smalln./spec.)	801
	nanny(ies) (hypoc./pej.)	797
	panty(ies) (spec./hypoc./pej.)	652

[54] Pejoration/hypocoristic(affection)/smallness

	starling(s) (spec./smalln.)	597
	Charley(s)|*Charlie(s)* (spec./hypoc./pej.)	541
	yearling(s) (smalln./spec.)	493
	pixie(s) (smalln./hypoc./pej.)	480
	Archie(s) (hypoc./spec.)	448
	dumpling(s) (pej./spec.)	545
	Collie(s)|*collie(s)* (spec./hypoc.)	533
	lassie(s)|*Lassie* (hypoc./spec)	510
	pixie(s) (smalln./hypoc./spec)	480
	stripling(s) (smalln./pej.)	409
	foundling(s) (smalln./pej./spec.)	376
	cockney(s)|*Cockney(s)* (spec./pej.)	345
	changeling(s) (pej./spec.)	311
	Total freq. of all dims. amb.	**50547**
Total size of corpus	406,232,024	109,939 (excl. amb).

Appendix B: Data used for the analyses of 4.3 and 4.4
Spanish lexemes from *CDE*:

No.	Diminutive lexeme	Raw	PMW	Diminutive meaning	Primitive, meaning & origin	Classification
1	PESETA(S)	3,234	78.2	'peseta'	*peso* 'weight'; peso (coin)'	Specialisation (Gooch 1967: 152)
2	SEÑORIITO/A(S)	2,097	50.7	'young gentleman; rich kid; lordling; young upper-class parasite; young, unmarried woman'.	*señor*, 'Mr.; man; God'	AMB: Smallness Pejoration
3	ORILLA(S)	2,097	50.7		*aura* 'aura	Non-diminutive
4	RODILLA(S)	2,085	50.4	'(Anat) knee' *rodilla* (diminutivo de rueda)	*rueda* 'wheel, circle'	Non-diminutive
5	BOLSILLO(S)	1,604	38.8	'pocket' (Gooch 1967: 90)	*Bolso* 'bag, handbag' (Gooch 1967: 90)	Specialisation (Gooch 1967: 8; 90)
6	SEMILLA(S)	999	24.2			Non-diminutive
7	CHIQUILLO/A(S)	801	19.4	'young boy, kid, lad'; 'young girl, las' (Gooch 1967: 83; 104)	*chico/a* 'boy; girl; small, little'	AMB: Smallness Pejoration
8	PASILLO(S)	786	19.0	'passage, corridor'	*paso* 'step' Gooch 93	Specialisation
9	TARJETA(S)	775	18.7	'small shield; card'	Borrowing from Old French *targette* 'small shield; card' (the small shield is where the currency is painted)	Non-diminutive
10	JUGUETE(S)	743	18,0	'toy(s)' or 'puppet, plaything (fig.)' (*WordReference* 2016)	*juego* 'game'	AMB: Specialisation

						Pejoration
11	CIGARRILLO(S)	743	18,0	'small cigar' (G91)	*cigarro* 'cigar'	Specialisation
12	GUERRILLA(S)	588	14,2	'guerrilla war(-fare), partisan war(-fare)', (as in 'small bodies of men acting independently'). Cf. *Guerrita* = young Guerra (surname) (Gooch 1967: 114)	*guerra* ‘war’	Specialisation
13	LADRILLO(S)	543	13,1	'clay brick, brick, galet, flagstone'	Lat. *later*'side', Ancient dim., from *ladre, from Lat. later,	Non-diminutive
14	MESETA(S)	462	11,2	'tableland, plateau' (Gooch 1967: 152)	*mesa* 'table'	Specialisation
15	BANQUETE(S)	455	11,0	‘banquet’ (Gocch 1967: 144)	*banco* 'bench'	Specialisation
16	ABANICO(S)	447	10,8	'fan; range; variety; derrick; (Caribe) points signal'	*abano* 'fan, punkah'	
17	CASITA(S)	447	10,8	'small house, (delightful) little house; cottage' (Gooch 57)	*casa* 'house'	Smallness Hypocoristic
18	CARRETA(S)	439	10,6	long narrow cart, waggon' (Gooch 151)	*carro* 'cart'	Specialisation
19	PAQUITA	438	10,6	Hypocoristic form of proper name	*Francisca* 'Frances'	Hypocoristic
20	GACETA	431	10,4			Non-diminutive
21	CAMPANILLA(S)	412	10,0	'small bell; blue-bell (flower); uvula'. Cf. *campanita* '(cute) little bell' (Gooch 1967: 111)	*campana* 'bell'	Specialisation
22	MALETA(S)	383	9,3	n. 'suitcase, bag'; adj. 'lazy'	*mala*'mailbag, mail, post'	Specialisation
23	TROMPETA(S)	363	8,8	'trompet; trumpeter (m.)' (Gooch 1967: 153)	*trompa* '[French] horn, instrument' (Gooch 1967: 153)	Specialisation

24	CUADRILLA(S)	348	8,4	‘bull-fghting team, team’ (Gooch 1967: 113)	*cuatro*'four'	Specialisation
25	CARETA(S)	343	8,3	*careta* 'mask, face shield' (Gooch 1967: 151)	*cara* 'face'	Specialisation
26	FESTÍN(ES)	338	8,2	‘feast, banquet’ (Gooch 1967: 75)	*fiesta* ‘party, festivity, holiday', from Fr. *festín*	AMB: Specialisation,
27	VENTANILLA(S)	325	7,9	'window (vehicle); counter position (bank, post office, etc.); nostril' (Gooch 119)	*ventana* 'window'	Specialisation
28	BOTÍN(ES)	305	7,4	'legging, ankle boot, sock'. (Primary meaning: 1SM [de guerra] booty, plunder; [de ladrón] loot).	*bota* 'boot' (Gooch 1967: 75)	AMB: Specialisation Smallnes Pejoration
29	MESITA(S)	289	7,0	'small table' (Gooch 1967: 59)	*mesa* 'table'	AMB: Smallness Hypocoristic
30	PUNTILLA(S)	278	6,7	sing. 'point lace; small nail; small dagger (bull-fighting); final stroke, last straw (fig.)' (Gooch 118 adv. (*de puntillas*) 'on tip-toe' (Gooch 1967: 124)	*punta*'tip, end'	Specialisation
31	BICICLETA(S)	264	6,4		Del fr. bicyclette, de bicy cle 'biciclo' y el suf. dim. -ette '-eta', por el menor tamaño de sus ruedas.	Smallness
32	GOLODRINA(S)	263	6,4	'swallow' (N. 'svale') (Gooch 1967: 81)	Del lat. *hirundo, -ĭnis.*	Non-diminutive
33	CAMISETA(S)	227	5,5	'vest' (Gooch 1967: 150)	*camisa* 'shirt'	Specialisation
34	MADRINA(S)	226	5,5	'godmother, sponsor' (Gooch 167:	*madre* 'mother'	Specialisation

				81)		
35	FRASQUITO(S)	220	5,3	Hypocoristic form of proper name	*Francisco*'Frank'	Hypocoristic (affection)
36	PASTILLA(S)	220	5,3	*pastilla* ‘ tablet; pastille; bar (of chocolate or soap) 1. f. Porción de pasta consistente, de forma, tamaño y usos variables, de uno u otro tamaño y forma. Pastilla de olor, de jabón.	*pasta* ‘paste’ Gooch 1967: 117)	Specialisation
37	BOLETÍN(ES)	220	5,3	'bulletin; (Univ) journal, review; (Escol) report'	*boleta* 'ticket; receipt', from it. *bollettino*, de *bolletta*. Italian	Specialsation
38	CAMIONETA(S)	215	5,2	'van' (G150)	camión 'lorry' Del fr. camionette, dim. d e camion 'camión'.	Specialication
39	PAJARITO(S)	206	5,0	'small bird, (sweet) little bird, "dicky" bird' (N. 'forbanna gjøk'?)		AMBIGUOUS: Smallness/hypocoristic (affection)/pejorative
40	NEBLINA(S)	194	4,7	'mist'	niebla 'fog'	Specialisation
41	RAMILLETE(S)	184	4,4	small bunch of flowers, bouquet, posy, nosegay	*ramo* 'branch; bunch of flowers'	Specialisation
42	MACETA(S)	174	4,2	1. 'flowerpot, plant pot'. 2. ' mallet, small hammer; stonecutter's hammer'	maza 1. 'mace, bat, stick, mallet' (N: 'klubbe, kølle') 2. a bore (person)	Specialisation
43	LIBRETA(S)	172	4,2	note-book-	libro 'book' (Gooch 1967: 152)	Specialisation
44	CORRILLO(S)	160	3,9	'‘small circle, group, gathering (e.g. for conversation or of spectators)' (G83)		AMB: Smalless/Pejorative
45	COQUETA	159	3,8	SF (=mueble) dressing table	*coqueta* from *coco* 1. f. rur. Ar. palmetazo.	Non-diminutive

					coqueta2 De coca5. 1. f. Ar. Panecillo de cierta hechura. coqueta3 V. coqueto.	
46	CAMARÍN(ES)	159	3,8	*camarín* '(theatre) dressing room (Gooch 71)	*cámara* 'chamber'	Specialisation
47	TORTILLA(S)	157	3,8	(G117): 'omelette'	from torta 'bun, cake, pastry'. Del dim. de torta.	Specialisation
48	BANQEUTA(S)	153	3,7	'stool, foot-stool'	banco 'bench'	Specialisation
49	BANQUILLO(S)	152	3,7	'small bench (navn på mindre benk); dock (Jur). (Gooch 1967: 89)	banco 'bench'	Specialisation
50	PANDILLA(S)	151	3,7	'gang'	(G117): panda 'band, faction',	Specialisation
51	CUARTILLA(S)	151	3,7	'sheet of writing paper' (Gooch 1967: 113)	cuarta 'forth part'	Specialisation
52	MULETA(S)	149	3,6	'crutch; muleta (bull-fight)' (Gooch 1967: 152)	from mula 'she-mule' muleto, ta 1. m. y f. Mulo pequeño, de poca edad o cerril. Mulo= muldyr	Specialisation
53	CEPILLO(S)	149	3,6	'brush; collection box (in church)'	cepo 'wooden stock'	Specialisation
54	CORNETA(S)	145	3,5	'bugle' (G151)	*cuerno* 'horn'	
55	PAPELETA(S)	144		(G152): 'slip of paper; ticket or receipt; pawn ticket; result paper (exam); trickey or unpleasant matter/situation'	papel ' paper'.	Spec.
56	CABECILLA(S)	144	3,5	(G123) 'ring-leader, rebel-leader, partisan leader'	*cabeza* 'head'	Spec.
57	HIJITA(S)	144	3,5			Ambiguous (Hypocoristic/smallness)

58	PATILLA(S)	141	3,4	'side-whisker(s), side-board(s); spectacle leg(s)	pata 'foot, leg'	Specialisation
59	TAQUILLA(S)	140	3,4	'box-office, booking office; window or grill (in banks, etc)'. (G118)	Del dim. de *taca*2 (uncommon) 'small cupboard'	Specialisation
60	BUHARDILLA(S)	139	3,4	'attic, garret, loft' (G111)	*buharda* 'skylight'	Specialisation
61	BARBILLA(S)	137	3,3	chin’	*barba* ‘beard’	Specialisation
62	PAJARILLO(S)	133	3,2	'(miserable) little bird' (G84)		Ambig. (Dim., often with pejorative implic.)
63	BARANDILLA(S)	132	3,2	(G110) 'railing, banister'	*baranda* 'rail'	Spec.
?	PIQUETE(S)	132	3,2	SEVERAL MEANINGS		
65	ENRIQUETA	129	3,1	Variante de *Enrique*, significa: Casa fuerte, rica, poderosa.	*Enrique*	Hypocoristic (affection)
?	SAINETE(S)	127	3,1			
67	BOQUETE(S) 121	121	2,9	‘narrow opening hole’. Also listed in (G144).	*boca* 'mouth'	Spec.
68	MARICA(S)	120	2,9	(G71): ‘pansy, sissy’ can aslo mean ‘magpie, but rarely used in this sense		Pejorative. Can also be used adjectively (Check hits in CDE?)
69	CALCETÍN(ES)	119	2,9	Gooch 75: ‘sock’	from *calceta* ‘hose’	AMB: Spec./smalln./pej
70	BORRICO(S)	119	2,9	'she-donkey; fool; sawhorse, sawbuck"	Del lat. tardío *burrīcus* 'c aballo pequeño', Eng. 'donkey; (Cono Sur) (hum) racehorse; (=perdedor en carrera) also-ran'	Not going use. Already diminutive in Latin when it came into Spanish.
71	BAYONETA(S)	119	2,9			
72	BOMBILLA(S)	118	2,9	Gooch 1967: 110): 'bulb (electric	Dim del *bomba* 'pump;	Specialisation

				light)' cf. *bombita* 'small light/bomb'	bomb'	
73	RATITO(S)	119	2,9	'short while' (Gooch 1967: 41)	*rato* 'a little while, a short time, a little bit'	AMB: Smallness Hypocoristic (Gooch 1967: 41).
74	MANTEQILLA(S)	116	2,8	'butter' (Gooch 1967: 115)	*manteca* 'lard; cooking fat; dripping'	
75	MALETÍN(ES)	115	2,8	'small or travelling suitcase' (Gooch 1967: 76)	*maleta* 'suitcase'	Specialisation (with a diminutive and somet. also pej. implications)

English lexemes from *COHA*:

	Diminutive lexeme	**Raw frequency**	**Type of dim.**	**Diminutive meaning**	**Primitive, meaning & origin**
1	CIGARETTE(S)	17968	Specialised	A small cigar made of a little finely-cut tobacco rolled up in thin paper, tobacco-leaf, or maize-husk.	cigar
2	DARLING(S)	11217	Hypocoristic (affection)	A person who is very dear to another; the object of a person's love; one dearly loved. Commonly used as a term of endearing address.	Old English *déorling*, *díerling*, derivative of *déor* dear n.1: see -ling suffix1.
3	DADDY(IES)	10525	Hypocorisic (affection)		
4	MAMMY(IES)/MOMMY(IES)/MUMMY/MAMY(IES)	2908	Hypocoristic (affection)		

5	PAMPHLET(S)	3859	Smallness		
6	AUNTIE(S)	2255	Hypocoristic (affection)		
7	BUNNY(IES)	2789	Hypocoristic (affection)		
8	GRANNY(IES)	2438	Hypocoristic (affection)		
9	BUGGY(IES)	2674	Smallness/Specialisation	A. 'A light one-horse (sometimes two-horse) vehicle, for one or two persons. B. Extended to various other vehicles or the like, as: (a) a perambulator (c) an automobile. Slang	
10	TROLLEY(IES)/TROLLY	1681	Specialisation	Locally applied to a low cart of various kinds, e.g. a costermonger's cart; at Yarmouth, a narrow cart or sledge adapted for the 'rows' or narrow alleys (row n.1 4b). Cf. troll n. 2. a. A low truck without sides or ends, esp. one with flanged wheels for running on a railway, or a track of rails in a factory, etc. Cf. bogie n. 1.	
11	HAMLET(S)	2026	Smallness	A group of houses or small village in the country; esp. a village without a church, included in the parish belonging to another village or a town. (In some of the United States, the official designation of an incorporated place smaller than a village.)	
12	PONY(IES)	6226	Smallness	orig. Sc. A small horse of any breed; spec. one not over a certain height (now usually 14.2 hands).	
13	TABLET(S)	2559	Specialisation		

14	JOCKEY(S)	1717	Hypocoristic (affection)	a. Diminutive or pet-form of Jock n.1; compare Jacky n.: originally Scots and northern English. ALSO: Transf. Of a thing. Cf. fellow n. 9, lad n.1, chap.A diminutive or familiar by-form of the name Jock or John, usually with the sense ‘little Jock, Jacky, Johnny’; hence, applicable (contemptuously) to any man of the common people (chiefly Sc.); also, a lad; an understrapper. (Cf. Jack n.1 2.)	
15	HIPPY(S)|HIPPIE(S)	2000	Hypocoristic (affection)	slang. A hipster; a person, usually exotically dressed, who is, or is taken to be, given to the use of hallucinogenic drugs; a beatnik.	*hip adj. + -y suffix6.*
16	COOKIE(S)	3055	AMB: Smallness/Specialisation	In Scotland the usual name for a baker's plain bun; in U.S. usually a small flat sweet cake (abiscuit in U.K.), but locally a name for small cakes of various form with or without sweetening. Also S. Afr. and Canad. **2.** SLANG: A woman; esp. an attractive girl. **3**.Usu. with modifying adjective expressing some positive personal quality: a person. Esp. in smart cookie, tough cookie.	probably < Dutch *koekje* /ˈkuːkjə/ diminutive of *koek* cake: this is apparently certain for U.S.; but for Scotland historical evidence has not been found.(Show Less)

17	BROWNIE(S)		Pejoration/hypocoristic(affection)/smallness	brownie1: A benevolent spirit or goblin, of shaggy appearance, supposed to haunt old houses, esp.farmhouses, in Scotland, and sometimes to perform useful household work while the family were asleep. 2. [< the colour of their uniform.] A member of the junior section of the organization known as the 'Girl Guides'. Also attrib. LEXEME 2: brownie | browny, n.2. 1. A sweet bread made with brown sugar and currants; currant bread. Austral. and N.Z. 2. A small square of rich, usu. chocolate, cake containing nuts. U.S. 3. An angler's name for: the trout.	Subst. use of *browny* adj. 'Inclining to brown'.< *brown* adj. + *-y* suffix1.
18	PUPPY(IES)	2339	Smallness/pejoration	†a. A small dog kept as a lady's pet or plaything; a lapdog. Obs. b. A young dog, esp. one that is less than a year old.c. In extended use: a young seal (or other young carnivore). Also: a young or small shark (cf. puppy shark n. at Compounds 2). Cf. pup n.1 3. rare. a. colloq. (freq. derogatory). A foolish, conceited, or impertinent young man; (also) a young person, esp. one who is inexperienced or naive	
19	SIBLING(S)				s*ie* 'Related by blood or descent; akin'. The derivatives from adjs. have the sense 'a person or thing that has the quality denoted by the adj**.**

20	SWEETIE(S)	819	Hypocoristic (affection)	1. orig. a. A sweetmeat, lollipop. Also, sweet cake or the like. b. attrib. 2. colloq. (orig. U.S.). A sweetheart, a lover; a lovable person. Also as a term of endearing address.	*sweet* adj. + -*ie* suffix.
21	LEAFLET(S)	1029	Smallness	A small leaf. b. Bot. One of the divisions of a compound leaf.c. popularly. A young leaf; rarely, a petal.3. A small-sized leaf of paper or a sheet folded into two or more leaves but not stitched, and containing printed matter, chiefly for gratuitous distribution.	: < leaf n.1 + -let suffix.
22	KINDLING(S)	801	smallness	VARIOUS SENSES: 1. 'The bringing forth of young'. 2. a. collect. 'A brood(=Progeny, offspring, young) or litter; progeny, issue'. b. sing. One of a brood or litter; a young animal'.	*kindle* v.2 a. trans. 'Of a female animal: To bring forth, give birth to (young)'. Also fig.
23	DEARIE(S)	683	Hypocoristic (affection)	Diminutive of dear. A. A little dear; a darling: a familiar term of amatory and conjugal endearment.	*< dear adj.1 + -ie suffix, -y suffix4.*
24	MISSY*	1229	Hypocoristic (affection)/pejoration	Used as an affectionate, playful, or (occas.) contemptuous form of address or mode of reference in speaking to or of a young girl or ‘miss’; (hence) a young girl.	*miss n.2 + -y suffix6.*
25	SAPLING(S)	1117	Smallness (a)/pejoration (b)	a. A young tree; esp. a young forest-tree with a trunk a few inches in diameter. 2. transf. A young or inexperienced person. 3. A young greyhound (see quots.).	*< sap n.1 + -ling suffix1. Compare sipling n.*
26	SEEDLING(S)	939	Smallness	A young plant developed from a seed, esp. one raised from seed as distinct from a slip, cutting, etc.	
27	PUSSY(IES)	1192	Hypocoristic (affection)	a. Chiefly colloq. A girl or woman exhibiting characteristics associated with a cat, esp. sweetness or amiability. Freq. used as a pet name or as a term of endearment. Cf. puss n.1 3,pussycat n. 3.	*puss n.1, -y suffix1.*

28	RIVULET(S)	1085	Smallness	a. A small river; a stream	*riveret n.;* *Latin rīvulus , -* *et suffix1.*
29	BOOKLET(S)	744	Smallness	1. A small book; spec. (in later use) a small, thin book with paper covers, containing information about a particular subject.	*book n., -let suffix.*
30	CADDIE(S)	687	Hypocoristic(affection)	a. A lad or man who waits about on the lookout for chance employment as a messenger, errand-boy, errand-porter, chair-man, odd-job-man, etc.; spec. a member of a corps ofcommissionaires in Edinburgh in the 18th c. (See also quot. 1883.) **Sc.**	French *cadet:* see *cadet* n.1 and *cad ee* n.
31	CHARLEY(S)|CHARLIE(S)	541	Hypocoristic (affection)	1. The name formerly given to a night-watchman. 2. A small triangular beard extending from the under lip, and ending in a point a little below the chin; well-known in the portraits of Charles I and his contemporaries. 3. Applied as a proper name to the fox. 4. pl. A woman's breasts. slang. 6. A fool, simpleton, esp. a proper, right Charley. slang. 7. (Also Mr. Charlie.) A white man. U.S. Black English slang. 8. U.S. Services' slang. The North Vietnamese and Vietcong; esp. a North Vietnamese or Vietcong soldier. 9. Used as **adj**.: Afraid, cowardly, esp. in phr. to turn Charlie. slang. Draft additions: slang (orig. U.S.). Cocaine.	a familiar variant of Charles.
32	ROULETTE(S)	522	Smallness	†1. A small wheel. Obs. rare. 2. Math. The path traced by a point on a curve as it rolls without slipping over another curve or a line. 3. a. A gambling game in which a ball is dropped on to a spinning horizontal wheel containing numbered and coloured compartments, with the players betting on which compartment the ball will come to rest in. Cf. roly-poly n. 2a. 5. A small tube around which hair is	*French roulette.*< *French roulette small wheel*

				rolled for curling; = pipe n.1 11. Now hist. and rare. ++	
33	COLLIE(S)	533	Hypocoristic (affection)	a. A Scottish shepherd's dog; a breed of sheepdogs remarkable for sagacity .b. fig. ‘One who follows another constantly or implicitly’ (Jamieson); cf. to dog, and Sc. follow-dog.	Origin uncertain: it has been conjectured to be the same word as *coaly* ‘the colour being originally black’; compare *colly* adj. C haucer has *Colle* as proper name of a dog, of which collie might possibly be dimin
34	PANTY(IES)	652	Hypocoristic (affection)	1. Men's trousers or shorts; boy's short trousers. Usu. in derogatory contexts. Now rare.2 a. Women's or girls' underpants; esp. short-legged or legless pants with an elasticated waist (now the usual sense). Also in extended use. b. slang. to get one's panties in a bunch (also wad, knot, etc.): to become agitated or annoyed. Cf. to get one's knickers in a twist at twist n.1 16e. II. In singular form 3. = sense 2a.	pants n., -y suffix6.< pant- (in pants n.) + -y suffix6. Compare pant n.3

35	BUDDY(IES)	2813	Hypocoristic (affection)	1. a. colloq. (orig. Caribbean, now chiefly U.S.). Used as a familiar form of address to a man or boy, sometimes one whose name is not known. b. colloq. (chiefly U.S.). A friend, a pal. 2. Caribbean and U.S. (chiefly south. and Midland) colloq. A brother. 3 a. Originally: either of a pair of miners working together in the same area of a mine. Later more generally: either of two individuals, each of whom takes responsibility for the other's safety or welfare, esp. during sporting activities. Also: either member of a pairing, esp. at a school or workplace, in which one member is assigned to provide support for the other. Cf. buddy system n. at Compounds.	1. *brother* n.Origin uncertain. Probably representing a regional pronunciation of brother n. (compare regional forms with medial u at that entry), with the ending perhaps influenced by association with -y suffix6. 2. The following earlier examples of *Buddy* as a term of endearment show a different formation (< *bud* n.1 3b + -*y* suffix6):
36	ARCHIE	448		Mil. slang (chiefly during or with reference to the First World War (1914–18)). Now hist. 1. Anti-aircraft guns collectively; anti-aircraft fire. Frequently personified.Chiefly applied by Allied soldiers to German guns. 2. As a count noun.a. A shell from an anti-aircraft gun. b. An anti-aircraft gun.	From a proper name. Etymon: proper name Archie.< Archie, pet form of the male forename Archibald (compare -y suffix6); for the motivation for the name see Archibald n
37	CASSETTE(S)	591		a. A casket. b. Photogr. (See quot. 1875). Also, a light-proof (cylindrical) container for a spool of film; a container for an X-ray plate or film.d. A closed container of magnetic tape with both supply and take-up spools, so designed that it needs merely to be inserted into a suitable tape recorder, computer, or video recorder to	French, diminutive of casse or caisse (com pare case n.1), < Italian cassetta.

				be ready for use. Cf. videocassette n. at video adj. Special uses.	
38	BULLY(IES)	1998		**†1.a**. A term of endearment and familiarity, orig. applied to either sex: sweetheart, darling. Later applied to men only, implying friendly admiration: good friend, fine fellow, 'gallant'. Often prefixed as a sort of title to the name or designation of the person addressed, as in Shakespeare, 'bully Bottom', 'bully doctor'. Obs. exc. arch. **b**. attrib., as in bully-boy. **2**. dial. Brother, companion, 'mate'. 3. a. A blustering 'gallant'; a bravo, hector, or 'swash-buckler'; now, esp. a tyrannical coward who makes himself a terror to the weak. **b**. A ruffian hired for purposes of violence or intimidation. arch. **4**. spec. The 'gallant' or protector of a prostitute; one who lives by protecting prostitutes.	Etymology obscure: possibly Dutch boel 'lover (of either sex)', also 'brother' (Verwijs & Verdam); compare Middle High German buole , modern German buhle 'lover', earlier also 'friend, kinsman'. Bailey 1721 has boolie 'beloved'

39	Dolly(ies)	1931	Hypocoristic/Pejoration/smallness **AMBIGUOUS**	1. A familiar pet-form of the name Dorothy (= doll n.1 1). 2.†a. A female pet or favourite. Obs. slang. b. A drab, slattern, useless woman. dial. or colloq. c. A girl or woman, esp. a young, attractive one. colloq. 3. a. A pet name for a child's doll. (Also treated as the personal name of a female doll.) b. Cricket (colloq.). (a) A donkey-drop (see donkey n. Compounds 2); (b) a very easy catch (cf. dolly adj. b). 4. Applied to various contrivances fancied to resemble a doll in some way. a. dial. A wooden appliance with two arms, and legs or feet, used to stir and twirl clothes in the wash-tub, called a dolly-tub; also called dolly-legs or dolly-stick, peggy, maiden.	doll n.1: see -y suffix1.

40	SHANY(IES)CHANTY|SHANTIE|SHANTEE	1430	Pejoration/diminutive **AMBIGUOUS**	1.a. Chiefly U.S. and Canada. A small, mean, roughly constructed dwelling; a cabin, a hut. c. = Canadian French chantier (see the etymology).2. Austral. and N.Z. A public-house, esp. unlicensed; a 'sly-grog shop'.	Probably corruptly < French chantier (see chantier n.) used in Canada in the senses: 'an establishment regularly organized in the forests in winter for the felling of trees; the head-quarters at which the woodcutters assemble after their day's work' (Franco-Canad., 1880, p. 38.
41	PIXIE(S)	480	Diminutive/(hypocoristic)	**(omtrent dss. nisse; adj skøyeraktig, ondskapsfull, ertelysten) 1**. In folklore and children's stories: a supernatural being with magical powers, typically portrayed as **small** and human-like in form, with pointed ears and a pointed hat. Also more generally (N. Amer.): **a fairy**. Also in extended use. **2. Chiefly Sc.** Short for pixie cap n. at Compounds or pixie hat n. at Compounds.	puck n.1, -sy suffix2.
42	STARLING(S)	597	Smallness	One or two names of birds have this suffix in Old English, asswertling ? some black bird (? < sweart black), stærling starling; here it may possibly have a diminutive force (see 2 below).https://vpn2.uio.no/+CSCO+1075676763 3A2F2F6A6A6A2E6272712E70627A++/view/Entry/10865 6#eid39234770 nonce-word: An inhabitant of a star.	Old English stærlinc , < stær stare n.1: see -ling suffix1.

43	JIMMIE(S)	385	Hypocoristic (affection)	Hypoc. For proper name James. Also: dismal Jimmy Jimmy Ducks	A male personal name, pet-form and familiar equivalent of the name James n.
44	LADDIE(S)	426	Hypocoristic	Formerly chiefly Sc. A young lad, a lad. (A term of endearment.)	lad n.1 + -ie suffix.
45	LASSIE(S)	510	Hypocoristic	Chiefly Sc. a. A lass, girl. b. = lass n. 1d. d. A female member of the Salvation Army.	lass n. + dimin. suffix -y suffix6)
46	YEARLING(S)	493	Smallness	**a.** An animal (esp. a sheep, calf, or foal) that is one year old; an animal in its second year. **Also occa**s. with reference to a child. **2. In pl**. Dried hops one year or more after harvesting, typically having a more mellow flavour than those in their first year. Now rare**.3.** A plant that is one year old; a plant in its second year. **4.** U.S. slang. a. At a military academy (esp. the United States Military Academy at West Point, N.Y.): a second-year cadet; a sophomore.b. A first-year college student; a freshman. Now esp. in sporting contexts.	year n. + -ling suffix1.
47	WEAKLING(S)	533	Pejoration	2. A person or animal that lacks physical strength, or is weak in health or constitution. 3. a. One who is weak in character or intellect. b. One who is a tiro or unskilled in (a subject).4. One who is weak in the faith or in spiritual attainments .5. appositive or as adj. Weak, feeble.	obs. ADJ BASE. weak adj. + -ling suffix1. Compare German weichling effeminate man.

48	NANNY(IES)		Ambiguous (hypocoristic/pejorartion)	a. A person, esp. a woman, employed by a family to look after a child or children; a children's nursemaid. Also as a familiar form of address. Cf. nan n.2 1, nana n.1 1 .b. A grandmother. Freq. as a form of address (used esp. by a child). Cf. nan n.2 2, nana n.1 2. 2. A person, institution, etc., considered to be unduly protective or interfering.	Probably < nan n.1 + -y suffix6, perhaps after Nanny, pet-form of the female forenames Anne and A gnes (see nan n.2; compare -y suffix6
49	DUMPLING(S+B87:G87)	545	Pejoration	a. A kind of pudding consisting of a mass of paste or dough, more or less globular in form, either plain and boiled, or enclosing fruit and boiled or baked. (Originally attributed to Norfolk.) b. transf. A pasty mass like a dumpling. 2. A dumpy animal or person, short and of rounded 1 (matlaging) melbolle dim?? 2 (matlaging, også apple dumpling) innbakt eple dim?? 3 (hverdagslig) trulte, liten tjukkas DIM!	probably < same source as dump **adj.**: see -ling suffix1.
50	STRIPLING(S)	409	AMBIGIOUS Smallness Pejoration	**1**. A youth, one just passing from boyhood to manhood. **2.** attrib. (chiefly appositive) passing into **adj**. **Check hits in COHA** Norsk: ung person, jypling(='liten narr, guttunge') ung, uerfaren (og viktig) fyr, grønnskolling)	Probably < strip n.1 (though that word is not recorded before the 15th cent.) + -ling suffix1.he etymological notion seems to be 'one who is slender as a strip', one whose figure is not yet filled out.
51	DROPLET(S)	433	Smallness	A minute drop. droplet infection, infection conveyed by fine droplets of mucus sprayed into the air when a person opens his mouth to speak, cough, etc.	let suffix.
52	FOUNDLING(S)	376	pejoration? PITY smallness??	a. A deserted infant whose parents are unknown, a child whom there is no one to claim. Also transf. b. fig. 2. the Foundling: the Foundling Hospital, London.	

53	ISLET(S)	546	Smallness	1. A little island, an eyot or ait. 2. transf. a. Something resembling an island in position; a small piece of land markedly differing in character from that by which it is surrounded, as a wooded eminence in a marsh or plain; any isolated tract or spot; = island n. 2a. Also = island n. 2d. b. An isolated piece of animal or vegetable tissue. islet (or Islet) of Langerhans	French islette, modern French îlette, diminutive of isle n.: see -et suffix1. See also islot n., isolet n
54	COCKNEY(S)	345	Pejoration	4. spec. a. A person born in the city of London: strictly, (according to Minsheu) 'one born within the sound of Bow Bells'. b. One of the 'Cockney school': see B. 2b.c. The dialect or accent of the London cockney or of those from the East End of London generally.B. adj. (orig. attrib. use of the noun). 1. Cockered, petted; effeminate; squeamish. 2a. Pertaining to or characteristic of the London Cockney.	Middle English coken-ey, -ay, apparently = coken of cocks + ey, ay (Old English æg) egg; lit. 'cocks' egg'.
55	HOMEY(S)|HOMIE(S)	306	Hypocoristic	HOMEY; colloq. Resembling or suggestive of home; home-like; having the feeling of home; homish. HOMIE: 1. N.Z. colloq. (freq. derogatory). An immigrant (usually a recent one) to New Zealand from Britain; any English (or British) person. Cf. Pommy n.2. slang (orig. and chiefly in African-American use; also among Hispanic Americans). A person from one's home town or neighbourhood; a member of one's peer group or gang; a homeboy or homegirl. In later use also: a member of the hip-hop subculture. Also as a form of address.	HOMEY< home n.1 + -y suffix1. For analogical spelling, compare bony, limy. HOMIE home n.1, -y suffix6.< home n.1 + -y suffix6. With later use in sense 2 compare homeboy n., homegirl n., and also homes n.
56	STREAMLET(S)	354	Smallness	A small stream; a brook, rill, or rivulet.	stream n. + -let suffix.
57	CIRCLET(S)	304	Smallness	**1**. A **small** circle (in various senses of that word). 2.a. spec. A ring or band (e.g. of precious metal or jewels) worn as an ornament, esp. on the head. **b**. gen. A ring, circular band, or **small hoop** of any kind.	French cerclet, diminutive of cercle ; subseq. influenced by English circle n., and probably by Italian circoletto : see -et suffix1

58	LEAFLET(S)	1029		**A small leaf**. **1**. **b**. Bot. One of the divisions of a compound leaf. **c.** popularly. **A young leaf**; rarely, a petal. **2**. **Physiol. and Zool**. An organ or part of an organ **resembling a small leaf**; spec. the thin flap of a valve in the heart or a blood vessel. **3**. **A small-sized leaf of paper or a sheet folded into two** or more leaves but not stitched, and containing printed matter, chiefly for gratuitous distribution.	leaf n.1 + -let suffix.
59	CHANGELING(S)	311	pejoration	**A. n. 1**. **One given to change;** a fickle or inconstant person; a waverer, turncoat, renegade. arch. **2**. A person or thing (surreptitiously) put in exchange for another. ? Obs. (exc. as in A. 3).**3. spec**. A **child** secretly substituted for another in infancy; esp. **a child (usually stupid or ugly**) supposed to have been left by fairies in exchange for one stolen. (In quot. 1600 applied to the child taken, not to that left.) =**(PEJ.) 4**. A **half-witted person,** idiot, imbecile. arch **=(pej.)**	change v. + - ling suffix1, diminutive suffix.G97F97C97:G97F97:G97
60	UNDERLING(S)	384		**A. n. 1. a**. One who is subject or subordinate to another; in later use esp. a subordinate agent or official, an understrapper. **b**. A branch, plant, etc., growing under, or less strongly than, another; a small or weakly plant, animal, or child. Now dial. 2. a. In predicative use, passing into adj.: Subject, subordinate (to a person, etc.). **b**. Similarly in attributive use. **B. adj**. 1. Undersized, small, weak. (Cf. underline adj.) **2**. Low-growing. **3.** Trivial, unimportant.	Early Middle English, < under adv. 3 + - ling suffix1.