

University of New Mexico
UNM Digital Repository

Linguistics ETDs

Electronic Theses and Dissertations

5-1-2014

"Hey, y'guys!": A diachronic usage-based approach to changes in American English address

Ben Sienicki

Follow this and additional works at: https://digitalrepository.unm.edu/ling_etds



Part of the [Linguistics Commons](#)

Recommended Citation

Sienicki, Ben. "Hey, y'guys!": A diachronic usage-based approach to changes in American English address." (2014).
https://digitalrepository.unm.edu/ling_etds/31

This Dissertation is brought to you for free and open access by the Electronic Theses and Dissertations at UNM Digital Repository. It has been accepted for inclusion in Linguistics ETDs by an authorized administrator of UNM Digital Repository. For more information, please contact disc@unm.edu.

Ben Sienicki

Candidate

Linguistics

Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Melissa Axelrod, Chairperson

Joan Bybee

Christian Koops

Joanne Scheibman

Catherine Travis

**“Hey, y’guys!”: A diachronic usage-based
approach to changes in American English address**

Ben Sienicki

MA, Applied Linguistics, University of Montana, 2001
BA, Philosophy, University of Montana, 1998

DISSERTATION

Submitted in partial fulfillment of
the requirements for the degree of:

PhD, Linguistics, May 2014

University of New Mexico
Albuquerque, New Mexico

**“Hey, y’guys!”: A diachronic usage-based
approach to changes in American English address**

Ben Sienicki

MA, Applied Linguistics, University of Montana, 2001

BA, Philosophy, University of Montana, 1998

PhD, Linguistics, University of New Mexico, May 2014

ABSTRACT

This dissertation adopts a functional, usage-based perspective on language to highlight key changes in American English address over the past century, especially the development of *you guys* and its expansion across second-person plural contexts. Based on data from the Corpus of Historical American English and the Corpus of Contemporary American English (among other corpora), the study tracks the increasing usage, gradual restructuring, semantic generalization, and shifting registers of *you guys*, including the interactions of those changes as the form has grammaticalized. This work offers an explanation, therefore, as to why *you guys* has been uniquely reshaped into a pronominal unit with non-masculine meanings in American English, while other appositive uses such as *you men* and *you fellows* have retained their structural and semantic properties with far greater fidelity.

Contents

1. Introduction	1
1.1 A brief history of English 2p pronouns.....	5
1.2 A functional usage-based approach to <i>you guys</i>	11
1.3 Hypotheses and research organization.....	16
2. Methods.....	18
2.1 The COHA data.....	18
2.2 The COCA data	20
3. The lexical basis of <i>you guys</i>	22
3.1 <i>You guys</i> 's entrance in the [PRO (AP) AppNP] construction.....	23
4. The role of frequency in the development of <i>you guys</i>	26
4.1 A comparison of <i>you</i> -appositives	28
4.1.1 The diachronic frequency of <i>you men</i> and <i>you fellows</i>	29
4.1.2 The frequency of <i>you all</i>	32
4.1.3 Summary and discussion.....	34
4.2 A frequency comparison of <i>you guys</i> and <i>you-p</i>	35
4.2.1 The frequency of <i>you-p</i> in the COHA.....	36
4.3 <i>You guys</i> 's higher frequency in speech: the COCA data	40
4.4 Reconciling corpus frequencies and survey rankings of 2p addressives	42
4.4.1 Corpus frequencies of contemporary 2p addressives	45
4.4.2 Summary and discussion.....	47
5. The role of chunking in the development of <i>you guys</i>	50
5.1 Changes in the adjacency of <i>you</i> and <i>guys</i>	51
5.1.1 The adjacency of <i>you</i> and <i>guys</i> in contemporary speech	55
5.1.2 Changes in adjacency among other <i>you</i> -appositives	57
5.1.3 Summary and discussion.....	62
5.2 The increasing association of <i>you</i> and <i>guys</i> in the appositive construction.....	63
5.2.1 Further evidence of association: the COCA spoken data.....	64
5.3 The univerbation of <i>you guys</i>	66
5.3.1 <i>You guys</i> versus other <i>you</i> -appositives in genitive constructions.....	66
5.3.2 <i>You guys</i> versus other <i>you</i> -appositives with contracted auxiliaries.....	70
5.3.3 A diachronic look at <i>you guys</i> 's univerbation	74
5.3.4 Orthography as a reflection of univerbation.....	75
5.4 Summary and discussion.....	77
5.4.1 Interdependence of chunking and frequency in <i>you guys</i> 's development	80
6. The semantic generalization of <i>you guys</i> and its social dynamics	82
6.1 The origin and conservation of <i>you guys</i> 's masculine semantics.....	84
6.1.1 The role of "tough guy" talk in <i>you guys</i> 's masculine semantics.....	88
6.2 <i>You guys</i> 's indefinite-inclusive semantics	90

6.3 <i>You guys</i> 's mixed-gender definite semantics	92
6.3.1 The role of intimacy in <i>you guys</i> 's early mixed-gender semantics	94
6.4 <i>You guys</i> 's feminine-exclusive semantics.....	99
6.5 Overview of <i>you guys</i> 's non-masculine definite semantics	102
6.5.1 Summary and discussion.....	104
6.6 Interdependence of semantic generalization and frequency in <i>you guys</i> 's development	106
6.7 Summary and discussion.....	111
6.7.1 <i>You guys</i> 's semantic change from a cross-linguistic perspective	113
7. Conclusions	115

“Hey, y’guys!”: A diachronic usage-based approach to changes in American English address

1. Introduction

The usage of an addressive pronoun is inherently a social act (Brown and Gilman 1960, Brown and Levinson 1987, Wales 1983, Mühlhäusler and Harré 1990, and Walker 2003). As address itself varies across the myriad discourse situations, regions, and interlocutors encountered from day to day, speakers adapt addressive expressions to different contexts of usage, and thus variability emerges within the class of addressives as a whole. Such is the case of the second-person plural (2p) addressives of American English, the multiplicity of which reflects the very diversity of American discourse and social interaction: *you*, *you all*, *y’all/ya’ll*, *you guys*, *yous(e)*, *yous(e) guys*, *yins/yinz*, *yuns*, *you people*, *you folks*, and numerous others. The primary goal of this dissertation is to trace the origin, development, and social history of one such form, *you guys*, which has been gaining traction in American English address.

By virtue of its relationship to *you*, the newer *you guys* addressive has fulfilled the range of morphosyntactic functions that personal pronouns can serve in English, from its incipience as a multi-part appositive construction to the present day. To illustrate this point, examples (1)-(12) compare earlier and later *you guys* uses as subjects, predicate complements (direct objects or other arguments), vocatives, and possessive determiners:

Subjects:

- (1) 1914: “Aw, say, *you guys* give me a pain!” (COHA, Zone)
- (2) 2000: “So, *you guys* both went to Laramie High?” (COHA, Laramie Project).

Predicate complements (direct objects):

- (3) 1930: “I won’t forget *you guys*” (COHA, Cimarron).

- (4) 2004: “Anyhow, I came down to say hello and invite *you guys* over for some games later” (COHA, Fat Pig).

Predicate complements (other):

- (5) 1920: “I thought *you guys* mighty cruel to put all that scare into a crowd in their condition...” (COHA, Eye Zeitoon).
- (6) 2006: “Did she give *you guys* any coffee? I’ll bet she didn’t” (COHA, Sleeper Novel).

Objects of preposition:

- (7) 1924: “The only difference between me and the rest of *you guys* is that I admit that I chase around with rats...” (COHA, Plastic Age).
- (8) 2003: “I’m glad that we’re all here with nowhere to go tonight, because I always wanted to play this thing with *you guys*” (COHA, Leslie Novel).

Vocatives:

- (9) 1923: “Give me the once-over, *you guys*. Do I look like a murderer?” (COHA, Adding Machine).
- (10) 2000: “Okay, so *you guys*, you and Russ go to the Fireside” (COHA, Laramie Project).

Possessive determiners:

- (11) 1949: “If Louie don’t come back it’s *you guys*’ fault” (COHA, Man with Golden Arm).
- (12) 2005: “*You guys*’ sister is getting married tonight” (COHA, Ploughshares).

The fact that the same set of grammatical roles has been maintained for *you guys* since its first wave of usage reveals its distinctive development with respect to *you*. *You*, once

functioning as an accusative/dative plural pronoun, later adopted nominative and singular functions (see Section 1.2). *You guys*'s more notable changes have occurred, by contrast, internally, or within its periphrastic structure and meaning.

I explain four interconnected changes in particular that have been integral in shaping *you guys*'s morphosyntax and semantics during the 20th century:

- (a) significant gains in token frequency, involving strong correlations with processes (b)-(d) below (Bybee and Thompson 2000);
- (b) gradual erosion of constituency structure by way of chunking (Bybee and Scheibman 1999, Bybee 2010);
- (c) loss of restrictions to masculinity via semantic generalization (Givón 1975, Bybee et al. 1994), including the adoption of indefinite inclusive, mixed-gender definite, feminine exclusive, and non-human animate meanings;
- (d) initial strengthening of lexical associations to contexts of social intimacy as newer non-masculine meanings emerged (i.e. pragmatic strengthening, Dahl 1985, Traugott 1988), followed by weakening of such associations through further semantic generalization.

As these changes are characteristic of grammaticalization more generally (Traugott and Heine 1991, Bybee et al. 1994, Bybee and Thompson 2000, Hopper and Traugott 2003), a diachronic usage-based approach to *you guys* incorporating grammaticalization theory is more suitable than prior approaches that treat *you guys* as a synchronic outcome of rule-application (Smith 1964, Postal 1966/1969, Delorme and Dougherty 1972) or ordinary word formation (Howe 1996). I develop the alternative view that *you guys* has

been influenced by the gradual processes of grammaticalization in American English since the turn of the century, becoming increasingly autonomous from the appositive construction in which it arose and on which prior approaches have been too narrowly based.

In light of (a)-(d) above, an investigation of *you guys* allows us to examine the relationship between grammaticalization and its subprocesses, including chunking, semantic generalization, and pragmatic strengthening. Because, as I show for *you guys*, these processes depend ultimately on domain-general cognition, statistical patterns in usage, and the social composition of the discourse, the present study illustrates the epiphenomenal nature of grammaticalization (Bybee 2010: §6.5) and the emergent quality of its outcomes (Hopper 1987). The story of *you guys* offers a glimpse at the manner in which numerous small-scale changes across behavioral domains—including non-linguistic ones—gradually accumulate to promote larger-scale grammatical outcomes, such as changes in American English address.

In clarifying the role of *you guys* in changes within American English address, this paper bridges work in the fields of diachronic linguistics, grammaticalization theory, functional usage-based grammar, corpus-based linguistics, and discourse analysis. The emergence of *you guys* and its role in 2p address in the United States has yet to be investigated from such an integrated perspective. In fact there have been very few substantive accounts of *you guys* from any theoretical background in the field. Among scholars who do mention the form (e.g. Lawson 1982, Jochnowitz 1983, Finegan and Rickford 2004, Murray and Simon 2008), it is most often given impressionistic or marginal treatment. Lawson 1982 and Jochnowitz 1983, for instance, provide but a few

anecdotal examples of *you guys* usage, while Finegan and Rickford 2004 and Murray and Simon 2008 offer only a brief remark concerning the form's existence as a 2p variant.

Arguably the most intensive treatment of *you guys* in the American English is given by Heyd 2010, who identifies social and situational constraints on *you guys* usage in episodes of the television show *Friends*. While her corpus-based approach overlaps methodologically with mine, she does not, however, concentrate on gradual change due to grammaticalization, being limited in this respect by the brief time-depth of her data, the *Friends* corpus. In addition, because this corpus is relatively small (590,000 words), represents only a handful of authors, and is based exclusively on fiction, her conclusion that *you guys* is an “emergent second-person plural pronoun” (2010: 34) remains largely speculative. My work not only anchors this claim more firmly in larger-scale, multi-genre diachronic corpora—and, wherever possible, authentic discourse—but also enhances our understanding of *you guys* as an instance of grammaticalization.

1.1 A brief history of English 2p pronouns

An account of *you guys* would lack sufficient historical context without first pausing to examine prior changes in English address. These changes have been well-studied in recent centuries (Kennedy 1915, Stidston 1917, Brown and Gilman 1960, Mulholland 1967, Traugott 1972, Baugh 1978, Wales 1983, Yang 1988, Mühlhäusler and Harré 1990, Walker 2003, Kytö 2004). Below I synthesize much of the existing research in an effort to better situate the contemporary usage of *you guys* and its related *you*-plural (*you-p*) form.

Throughout the Old English (OE) period (~500s-1000s ACE), the second-person pronominal system involved three distinctions across number—singular, dual, and plural—and three across case. Refer to Table 1:

	Singular	Dual	Plural
Nominative	ƿu	ġit	ġe
Accusative or Dative	ƿe	inc	ēow
Genitive	ƿīn	incer	ēower

Table 1: The second-person pronominal system of OE (adapted from Campbell 1959: 288)

By the Middle English (ME) period (~1100s-1400s ACE), the dual forms of Table 1 had fallen out of usage (Baugh 1978: 58-59), and only a two-way distinction across number remained, shown in Table 2:

	Singular	Plural
Nominative	þou/thou	ȝe/ye
Accusative or Dative	þee/thee	ȝow/you
Genitive	þy(n)/thy(n)	ȝower/your

Table 2: The second-person pronominal system of ME

The slash mark in Table 2 is used to distinguish the predominant orthographic variants of second-person pronouns in ME. The forms on the left represent older variants than those on the right.¹

The second-person pronouns of Table 2 underwent yet another transition in late ME (~1400s), which continued roughly a century into the Early Modern English (EModE) period (~1500s-early 1700s ACE). In the 1400s and 1500s, the plural forms became increasingly used in formal registers, in the address, for one, of superiors in social class (Kennedy 1915, Stidston 1917). The social rank meaning expressed by *ye/you/your* was then gradually generalized to contexts of politeness, even when

¹ Other orthographic variants exist as well. In the Penn corpus, for instance, ‘g’ is sometimes substituted for ‘y’ among the plural forms, creating structural indistinctness at times between the past participle marker *ge-* and the nominative plural form *ge*. All known orthographic variants have been included in Table 2.

interlocutors were of equal social class—what Baugh 1978: 242 labels the ‘courtesy reading’.

Once the meaning of the plural forms in Table 2 were associated with formality and politeness, the forms began to be used further in singular contexts to avoid the inference that the speaker was—by otherwise using an informal *thou* form—acting condescendingly or impolitely toward the interlocutor. Historical texts from this period abound with the socially stigmatized *thou* meaning. In the trial of Sir Walter Raleigh in 1603, for example, the prosecutor at one point recruited *thou* forms to disparage the defendant: “I thou thee, thou traitor!” (cited in Mencken 1936: 450). Shakespeare used *thou* forms similarly in *Twelfth Night*. In one scene, Sir Toby Belch and Sir Andrew Aguecheeck attempt to incite Viola, who has disguised herself as a man, to fight them by addressing her with *thou* (ibid.). Such a scenario involving the stigmatization of *thou* forms likely had the effect of limiting their usage, while conversely *ye*, *you*, and *your* became more common as they generalized to singular (polite) contexts (Wales 1983: 117, also see Walker 2000, 2003). This process eventually led to the erosion of the restriction to plural contexts for *ye/you/your*.

Consequently by late ME, most speakers had begun to associate the singular forms of Table 2 with contexts of informality, to be used “among familiars [e.g. to express solidarity] and in addressing children or persons of inferior rank” (Baugh 1978: 242, see also Brown and Gilman 1960, Brown and Levinson 1987). Other scholars argue that this view oversimplifies matters (e.g. Wales 1983, Mühlhäusler and Harré 1990, and Walker 2003), as it overlooks additional motivations for a speaker’s selection of a second-person addressive in the late ME and EModE periods. They point out that the

thou forms were also used in private discourse, in the expression of heightened emotion (e.g. anger), and in the reinforcement of intimacy between friends or family members. The *ye* forms, by contrast, tended to be used more in public, in contexts that were unemotional, and in situations in which respect rather than intimacy characterized the social relation.

Wales 1983 and her colleagues further stress that such motivations were tendencies rather than absolutes, indicating that there was much more variation in late ME and EModE than Brown and Gilman 1960 and their associates suggest. The usage of *thou* to express a shared belief in egalitarianism, for instance, was common in certain religious groups such as the Society of Friends (Quakers), a point which Brown himself later concedes (1986). This meaning of *thou* is quite distinct from the one found in the trial of Sir Walter Raleigh or *Twelfth Night* and has in fact persisted into the present day in offshoot religious groups and other (secular) dialect communities (Wales 1983, 2004). Thus the picture of the second-person pronominal system of late ME and EModE is far “messier” than that originally depicted by Brown and Gilman. Refer to Table 3:

	Singular	Singular or Plural
	<u>Addressive contexts:</u> social inferiors, social equals (lower classes), private, informal/intimate, heightened emotion	<u>Addressive contexts:</u> social superiors, social equals (upper classes), public, formal/neutral, respect
Nominative	thou	ye
Accusative or Dative	thee	you
Genitive	thy(n)	your

Table 3: The second-person pronominal system of late ME/EModE (adapted from Wales 1983: 116)

Brown and Gilman’s underestimation of the contextual factors underlying *thou-ye* usage is arguably based on the orientation of their research program toward sociolinguistic

universals; being qualitative, their methodology also relied heavily on extrapolation from a limited set of examples and prior scholarship, rather than on corpora (in contrast, for instance, to Walker 2000, 2003).

Given the expansion of the erstwhile plural forms to singular contexts, these forms gain in frequency in late ME. To illustrate this point, a search of *ye* (including its orthographic variants) in the early ME portion of the Penn corpus (~1100s-1200s) yields a normalized frequency of 13.8 per 10,000 words, but by late ME (~1300s-1400s) the form increases in usage by 25%, occurring 17.2 times per 10,000 words. Due to the particularization of discourse contexts for the *thou* forms (Table 3), they in turn decrease in frequency in late ME. *Thou*'s rate of occurrence, for example, drops from 49.3 per 10,000 words in early ME to 32.9 per 10,000 words in late ME, reflecting a 33% decline.

The endpoint of changes in 2p address during late ME appears in fuller view in the EModE period, as the two most frequent second-person pronouns (*you* and *your*) become even more frequent, while the other four forms decline or stabilize in usage. Table 4 compares the rate of occurrence of second-person pronouns in early EModE (~1500s) and late EModE (~1600s-early 1700s), listed in ascending order of frequency in the latter period:

Form	Early EModE (~1500s)		Late EModE (~1600s-early 1700s)	
	Token Frequency	Frequency (per 10,000 words)	Token Frequency	Frequency (per 10,000 words), late EModE
<i>thy(n)</i>	575	11.7	124	2.5
<i>thee</i>	308	6.2	156	7.4
<i>ye</i>	570	11.6	191	9.1
<i>thou</i>	523	10.6	288	13.7
<i>your</i>	628	12.7	388	18.5
<i>you</i>	655	13.3	500	23.9

Table 4: The frequency of second-person pronouns in early versus late EModE, orthographic variants included (Penn corpus, Early EModE: 493,218 words total, Late EModE: 209,564 words total)

Throughout the EModE period, *you* and *your* remain the most frequent second-person pronominals, increasing noticeably in usage in late EModE.

In EModE, *you* forms were slowly replacing *thou* forms in addressive contexts involving social inferiors, social equals (among lower classes), private confidants, and so on, in which *thou* would have been previously used (Wales 1983, Mühlhäusler and Harré 1990, and Walker 2003). This process gave rise to the two-form *you/your* system of Modern English seen below.

	Singular or Plural
Nominative, Accusative, or Dative	<i>you</i>
Genitive	<i>your</i>

Table 5: The pronominal system of Modern English

In the Modern English period, *you* became even more polysemous across case as it began to perform the nominative functions of *thou* and *ye*, in addition to its older accusative and dative functions. Ostensibly from Table 5, it appears that fewer forms were at a speaker's disposal in 2p address in this period. Such a superficial interpretation of the data underestimates the actual variability of addressives, however, which also included periphrastic structures and regionalisms.

Since Modern English, an assortment of other variants have peppered 2p addressive contexts in the United States alongside *you/your: you all, y'all/ya'll, you guys, yous(e), yous(e) guys, yins/yinz, yuns, you people, you folks*, and various others. While the development of many such forms has been investigated thoroughly over the years (Axley 1926/1927, Hills 1926/1927, Vaux 1999, Tillery et al. 2000, Richardson 1984, Lipsky 1993, Maynor 1996, 2000, Wales 1996, to appear, Hickey 2004, Vaux and Jøhndal 2009), *you guys*, for one, has yet to be adequately addressed; this fact is ironic given *you guys*'s increasing impact on American English address and its interesting, somewhat controversial semantic history. The present work responds to this need by investigating the functional usage-based processes of language change that have contributed to *you guys*'s rise in American English by gradually modifying its structural, semantic, and pragmatic properties.

1.2 A functional usage-based approach to *you guys*

In this section I outline the main tenets of functional usage-based linguistics (e.g. Givón 1973, Hopper and Thompson 1984, Bybee 1985, DuBois 1985) to show why it is optimal for examining the development of *you guys* in American English.

Usage-based theory equates grammar with the cognitive organization of one's experience with language (Bybee 2006). This notion stands in stark contrast to generative models of language (Chomsky 1957, 1965), in which grammar, or 'competence', is thought to arise independently from usage, or 'performance', and consequently one's experience with language falls outside the scope of grammatical research. In one of its most recent strains, functional usage-based linguistics incorporates aspects of grammaticalization theory, construction grammar, typology, cognitive linguistics,

exemplar models of categorization, variationism, and emergentism or complex adaptive systems research, summarized in Table 6:

Research area	Link to usage-based theory	Selected references
Cognitive linguistics	Linguistic categories result from the interaction of conceptualization, or general cognition, and our experience with language.	Lakoff and Johnson 1980, Lakoff 1987, Langacker 1987, 1991, Talmy 1988, 2000, Croft and Cruse 2004
Construction grammar	Constructions, or multi-word sequences ranging from more to less fixed in their structure and meaning, are foundational to the cognitive organization of language. Lexical categorization depends on a speaker's assessment of this level of fixity and on the particular function of a construction at the moment of usage—to the extent that this function is prototypical of a larger class of stored items.	Fillmore et al. 1988, 1994, Goldberg 1995, 2003, Croft 2001 (also see Langacker 1987)
Discourse-based grammar	Morphosyntactic structure is rooted in the immediate discourse needs of speakers and in the structure of the discourse itself.	Hopper and Thompson 1980, 1984, Thompson and Mulac 1991
Emergentism, Complex adaptive systems	Grammatical outcomes arise indirectly from a variety of local processes (e.g. individual usage events, repetition of a form in a particular construction, general-cognitive mechanisms, social constraints), rather than being determined directly by an innate, domain-specific grammar module.	Lindblom et al. 1984, Hopper 1987, Larsen-Freeman 1997, Bates and Goodman 1999, Elman 1999, MacWhinney 1999, Bybee 2010
Exemplar models	Each repeated experience with language has an impact on its storage and cognitive organization; grammar is shaped by associations between similar tokens of experience, or exemplars.	Nosofsky 1988, Johnson 1997, Pierrehumbert 2001, 2002, Bybee 2001, 2002, 2006
Grammaticalization theory	Lexical items develop into grammatical forms through usage, or grammatical forms develop new grammatical functions through usage.	Givón 1973, Lehmann 1982, Traugott 1982, Heine and Reh 1984, Bybee 1985
Typology	Structural regularities across languages depend on frequency of usage and common semantic sources for grammatical constructions.	Greenberg 1963, 1966, 1978, Croft 1990/2003; Heine and Kuteva 2002
Variationism	Gradience in linguistic categorization, which is influenced by frequency of usage and its distribution across pragmatic contexts, promotes variability in a particular class of lexical items.	Labov 1966, Sankoff 1980, Chambers 1995/2003, Chambers, Trudgill, and Schilling-Estes 2002, Tagliamonte 2012

Table 6: Core components of usage-based theory (alphabetical order)

Bybee 2010, for example, merges the core components of usage-based theory to identify the domain-general mechanisms of language change whereby linguistic structure emerges:

Since all patterns of linguistic structure have an evolutionary history, part of the explanation for why languages have particular structures must involve reference to how these structures arose. (10)

To this end, Bybee 2010 explores numerous non-linguistic processes that are central in language change, including chunking (Miller 1956, Newell 1990), habituation (Haiman 1994), and pragmatic strengthening (Dahl 1985, Traugott 1988). Below I examine how each process unfolds in language more generally. In subsequent chapters, I look into how they have interacted in the development of *you guys* in particular.

One domain-general mechanism that plays a formative role in language change is chunking. In this process, items that are repeatedly adjacent in experience become more strongly associated in memory, thus gradually forming a processing unit. Chess experts, for instance, have been known to use the strategy of grouping multiple moves into a single causal chain based on previous successes, “thinking ahead,” as it were, by thinking holistically (Chase and Simon 1973, Gobet, Retschitzki, and de Voogt 2004). In language, chunking has been linked to the emergence of morphosyntax (Bybee and Scheibman 1999, Bybee and Thompson 2000), or to changes in the constituency structure of sequential items in speech.

The reconfiguration of the English *be going to* construction into the processing unit illustrates chunking well. As *going* and *to* became increasingly sequential in speech, their constituency boundary began to erode, and gradually the *gonna* morpheme took shape. The chunk that emerged from this process has consequently become more autonomous from *go*'s earlier representations as a motion and purpose verb, as in utterances like “I'm *gonna* stay put” (lacking a motion meaning) and “They're all *gonna*

get lost if they don't use a map" (lacking a purpose meaning). Thus an important outcome of chunking is the weakening of cognitive associations to older formal representations of sequential items as separate constituents of a construction.

Another non-linguistic mechanism that is relevant in language change is habituation (Haiman 1994), which is likewise a correlate of frequent repetition. In habituation, the association between a particular behavior and its repeated context steadily erodes in memory, as the organism becomes more desensitized to the context. For example, a stray dog that has been taken in may growl and bear its teeth when it is initially presented with food by a new owner, but on subsequent feedings may only bear its teeth; in time, or through repetition of this scenario, it may perform neither act, simply taking the food. Such a change reflects a weakening of cognitive associations between the feeding action and contexts of potential aggression. Growling and teeth-bearing may in turn be recruited in a playful game of tug-of-war with the new owner, as the behavior generalizes beyond potentially aggressive contexts.

According to Bybee et al. 1994: §1.3, habituation undergirds functional changes in language as well, or semantic generalization, in which semantic content is lost as an item is frequently repeated in speech. Semantic generalization has occurred, for instance, in the weakening of lexical connections to *go*'s motion meaning in the *be going to* purpose construction, and in the weakening of lexical connections to *go*'s purpose meaning in the future construction with *gonna*. These changes reflect the gradual desensitization of speakers to *go*'s motion contexts as it became more frequently associated with those of purpose and futurity.

The development of *gonna* further illustrates the interaction of chunking and habituation in language change. On the one hand, chunking has contributed to the increasing dissociation of *gonna* from its prior formal representation as a three-part *be going to* phrase. Habituation, on the other hand, has led to *gonna*'s divorce from its previous functional manifestations, or from *go*'s earlier motion and purpose meanings.

While both chunking and habituation initiate reduction in language, a distinct process of language change leads conversely to the enrichment of an item's lexical representation. This process is known as pragmatic strengthening (Dahl 1985, Traugott 1988), otherwise referred to as the "enrichment" (Hopper and Traugott 2003). In pragmatic strengthening, recurring contextual conditions under which a form appears in discourse are repeatedly associated with the form's meaning, to the extent that the inference becomes lexically stored as part of the meaning. Continuing the example from above, as *go*'s earlier usage in purpose constructions was often paired with contexts wherein a future meaning could be inferred (e.g. "He is going to town to buy goods"), futurity later became part of *go*'s meaning. *Go*'s loss of restrictions to a motion meaning therefore coincided with the adoption of novel semantic properties stemming from implication and inference, or more generally from association (i.e. based on Aristotle's 'law of contiguity'). Insofar as inferencing and associative processes are non-linguistic in nature, enrichment represents an additional domain-general process underlying language change.

Because repetition is operative in chunking, habituation, and association alike (Bybee and Thompson 2000, Bybee 2010), measurements of frequency of usage are also vital in accounting for language change. One such measure is token frequency, or the

count of instances of a particular linguistic item in a text or corpus. The present study therefore centers on the relationship between the domain-general mechanisms of grammaticalization and token frequency at various stages in the development of *you guys*.

1.3 Hypotheses and research organization

If the development of *you guys* can be traced alongside increases in token frequency, gradual weakening of morphosyntactic constituency, enrichment, and semantic generalization, then a strong argument can be made that it has grammaticalized, as these are predictable changes in grammaticalization (Dahl 1985, Traugott 1988, Traugott and Heine 1991, Bybee et al. 1994, Bybee and Thompson 2000, Hopper and Traugott 2003). If, moreover, it can be shown that *you guys*'s formal and functional properties have stemmed from a diachronic process, then the form's traditional explanation as the synchronic output of generative determiner or appositive syntax (Smith 1964, Postal 1966/1969, Delorme and Dougherty 1972) turns out to be inadequate, if not altogether misleading. More recent characterizations of *you guys* as a "lexical compound" (Howe 1996) or "lexicalized form" (Wales 2004, De Vogelaer 2007) similarly disregard the role of grammaticalization in the form's development by emphasizing its lexical rather than grammatical character.

To substantiate the view that *you guys* has grammaticalized in American English and that its prior treatments are therefore limited, this paper adopts the diachronic usage-based approach outlined in Section 1.2, highlighting the cumulative impact of repetition (Chapter 4), chunking (Chapter 5), enrichment/association (Chapter 6), and semantic generalization (Chapter 6) on the form's development. Considered together, these

chapters add further weight to the usage-based principle that our experiences with language inform its structure.

Chapter 6 additionally argues that newer 2p addressives such as *you guys* have not merely arisen to “repair” or “fill a gap in” English’s pronominal paradigm, which has become a consistent current in the literature (Mencken 1936, Trudgill and Chambers 1991, Maynor 2000, Hickey 2003, Quinn 2009). Trudgill and Chambers 1991: 8 propose, for instance:

In many dialects of English around the world, the historical loss of the second person singular/plural distinction that went with the loss of *thou/thee* has been repaired by the introduction of new second person plural pronouns, such as *youse*, which is found in North America, Australia, Scotland, England, and especially Ireland.

Not only is such a view teleological and consequently dubious as an explanation for language change, but it further neglects the fact that nearly 20% of the world’s languages tolerate number gaps in their pronominal paradigms, a point advanced by Croft 2000. Adding to his critique, I argue that *you guys* cannot be explained away by its plural-marking function, having served other social functions as well at key turning points in its history.

2. Methods

To meet the objectives specified in Chapter 1, the principal method of the dissertation is corpus-based data collection and analysis. In particular, the Corpus of Historical American English (COHA), the Corpus of Contemporary American English (COCA) are utilized to track the usage and grammatical development of *you guys*. Importantly, the latter resource allows for measurements of naturalistic discourse, which are critical in building a maximally valid usage-based account of grammatical phenomena.

2.1 The COHA data

The COHA (406.2 million words) is particularly useful for tracing the full range of diachronic changes in *you guys*'s usage, as the database spans the previous two centuries from 1810 to 2009. Thus for the present study the COHA's advantage lies in its comprehensive time depth. Its disadvantage, in turn, is its general dearth of naturalistic discourse, given the scarcity of early recordings of spoken American English. This shortcoming is mitigated against, however, by adding further data from the spoken section of the COCA (Section 2.2).

As for the written texts that comprise the COHA, they are almost equally distributed across fiction (51%) and non-fiction (49%), occupying roughly comparable portions of the corpus from one period to next:

Period	Word Count					% Fiction
	Fiction	Magazines	Newspapers	Non-fiction books	Total	
1810s	641,164	88,316	0	451,542	1,181,022	54
1820s	3,751,204	1,714,789	0	1,461,012	6,927,005	54
1830s	7,590,350	3,145,575	0	3,038,062	13,773,987	55
1840s	8,850,886	3,554,534	0	3,641,434	16,046,854	55
1850s	9,094,346	4,220,558	0	3,178,922	16,493,826	55
1860s	9,450,562	4,437,941	262,198	2,974,401	17,125,102	55
1870s	10,291,968	4,452,192	1,030,560	2,835,440	18,610,160	55
1880s	11,215,065	4,481,568	1,355,456	3,820,766	20,872,855	54
1890s	11,212,219	4,679,486	1,383,948	3,907,730	21,183,383	53
1900s	12,029,439	5,062,650	1,433,576	4,015,567	22,541,232	53
1910s	11,935,701	5,694,710	1,489,942	3,534,899	22,655,252	53
1920s	12,539,681	5,841,678	3,552,699	3,698,353	25,632,411	49
1930s	11,876,996	5,910,095	3,545,527	3,080,629	24,413,247	49
1940s	11,946,743	5,644,216	3,497,509	3,056,010	24,144,478	49
1950s	11,986,437	5,796,823	3,522,545	3,092,375	24,398,180	49
1960s	11,578,880	5,803,276	3,404,244	3,141,582	23,927,982	48
1970s	11,626,911	5,755,537	3,383,924	3,002,933	23,769,305	49
1980s	12,152,603	5,804,320	4,113,254	3,108,775	25,178,952	48
1990s	13,272,162	7,440,305	4,060,570	3,104,303	27,877,340	48
2000s	14,590,078	7,678,830	4,088,704	3,121,839	29,479,451	49
Total	207,633,395	97,207,399	40,124,656	61,266,574	406,232,024	51 avg.
%	51	24	10	15	100	

Table 7: COHA overview, by period and genre

The balanced distribution of the COHA across fiction and non-fiction aids in controlling for genre-bias in measurements of *you guys*'s frequency, structure, and semantics in subsequent chapters, though this bias cannot be controlled for exhaustively due to potential differences across sub-genres (e.g. magazines versus newspapers).

In Chapter 4, I provide counts for all available *you guys* tokens in the COHA, arranging the results incrementally by decade, to show that *you guys* has increased significantly in usage over time. I additionally test for a positive and statistically significant correlation between the frequency of *you guys* and its decade of usage to confirm that the relationship between these variables is meaningful, or non-random.

To further demonstrate that *you guys* has become increasingly chunked, I track changes in the unithood of *you* and *guys* in the COHA and elsewhere (Chapter 5). I test

for correlations between these changes and *you guys*'s frequency from Chapter 4, in order to relate the form's restructuring to its usage and thereby illustrate the usage-based principle that our experience with language impacts its structure.

In Chapter 6, I track the meanings and addressive contexts of *you guys* across relevant periods of the COHA, including the first attestation of each novel meaning. A discourse analysis in this chapter identifies particular addressive contexts that were critical in *you guys*'s development of newer meanings. Ultimately I determine whether a positive correlation exists between the proportion of newer meanings (i.e. non-masculine ones) and the form's frequency over time, establishing the interconnection of these variables.

By finally relating *you guys*'s frequency gains (Chapter 4) to formal and functional changes throughout the 20th century (Chapters 5 and 6), I support the argument that *you guys* has been molded by usage-based processes of grammaticalization.

2.2 The COCA data

The COCA (464 million words) offers a diachronic perspective on *you guys* as well, yet with a much more limited time depth of only twenty two years (1990-2012). The relative advantage of the COCA, however, is its inclusion of discourse-in-use (spoken section). Consequently I employ the COCA wherever feasible to affirm and expound upon results deriving from the COHA, with the caveat that the COCA spoken section offers a narrow range of discourse situations given its basis in television news, talk, and interview shows.

The COCA is similarly balanced across communicative styles. With the difference across its respective portions being no greater than 1.3% (Table 8), the effect of style-bias can be largely factored out of the COCA results.

Period	Word Count					
	Spoken	Fiction	Magazine	Newspaper	Academic	Total
1990-1994	21,967,915	20,258,031	21,269,305	20,441,781	20,062,098	103,999,130
1995-1999	21,285,102	19,499,437	21,740,268	20,440,794	20,481,591	103,447,192
2000-2004	21,467,138	19,712,300	21,491,785	20,294,001	19,975,477	102,940,701
2005-2009	20,188,338	20,484,607	20,854,138	20,153,775	20,345,999	102,026,857
2010-2012*	10,477,179	10,389,759	10,209,210	10,350,615	10,179,613	51,606,376
Total	95,385,672	90,344,134	95,564,706	91,680,966	91,044,778	464,020,256
Percent	20.5	19.5	20.6	19.8	19.6	100.0

Table 8: COCA overview, by period and style (*=January-June 2012, based on date of last update)

At this stage it is important to note that when comparisons across corpus sections or corpora are made, I provide normalized token frequencies in addition to raw token frequencies—either in the text or in tables/figures—to avoid statistical biases related to sample size.

Based on the corpora thus far detailed, subsequent portions of the paper provide evidence that *you guys* has formed a pronominal unit with increasingly generalized meanings and pragmatic functions in 2p address. After first discussing the lexical basis of *you guys* (Chapter 3), I turn to the role of frequency in its development, followed by the roles of chunking and semantic generalization, including the complex interplay of enrichment and semantic generalization in the development of *you guys*'s novel meanings.

3. The lexical basis of *you guys*

In view of *you guys*'s previous treatment as lexical, or sub-grammatical, structure (Section 1.3), a principle concern of the current work is to determine the extent to which the form has become autonomous from its more lexical basis. It is therefore necessary to establish the fact that *you guys* has acquired a distinct set of formal and functional properties compared to the construction from which it derives.

The historical basis of *you guys* is an appositive construction in which a plural nominal marked for number renames, or designates, the set of referents noted by a preceding plural personal pronoun, in this case *you* (Postal 1966/1969, Delorme and Dougherty 1972). The appositive construction is a referring expression, in other words, with the deictic function of profiling a particular group of addressees in relation to the speaker. Examples include *you/we students* or *you/we Americans* in more standard speech varieties, and *us/them students* or *us/them Americans* in less standard varieties. The construction also entails an optional slot for an intervening adjectival phrase, as in *you lucky winners* or *you mean and nasty people*. Thus it is symbolized as [PRO (AP) AppNP] in subsequent chapters of the paper. In Chapters 4-6, I account for *you guys* as a special instance of this construction in which *you* and *guys* have become increasingly sequential, structurally and semantically reduced, and at pivotal stages in *you guys*'s semantic history, pragmatically enriched. Such changes in the erstwhile *you guys* appositive use epitomize the principle of diachronic usage-based linguistics that grammaticalization always occurs within particular constructions (Bybee et al. 1994, Traugott 2003), and consequently that constructions are primary loci of grammatical change.

Insofar as the [PRO (AP) AppNP] construction incorporates a pronoun, *you guys* is partially grammatical from the outset of its development. The [PRO (AP) AppNP] construction differs from more intensively grammaticalized personal pronouns like *you* and *we*, however, due to its greater structural “bulk” (Givón 1975)—being a multi-part phrase that incorporates other adjectival and nominal phrases. In Chapter 5, I argue that this trait of the construction renders it distinct from the *you guys* pronominal unit, which has gradually lost its AP slot and therefore structurally reduced. I conclude that the pronominal *you guys* unit has diverged from the [PRO (AP) AppNP] construction more generally, or that *you guys* has undergone a formal split in its representation (Heine and Reh 1984: 57-9).

3.1 *You guys*’s entrance in the [PRO (AP) AppNP] construction

The first documented usage of *you guys* in the OED is from 1896, occurring in the novella *Artie*, by the Chicago writer George Ade. In one scene of the novella, Ade’s character Mrs. Morton has convinced three men—Mr. Blanchard, Miller, and Artie—to attend a church fundraiser about which they are extremely unenthusiastic. After she leaves the room, the men bet cigars that the others will not make it, aware of their mutual reluctance to go. Near the end of this scene, Artie comments:

- (13) You’d better make it chewin’ gum. Next thing you’ll be bettin’ real money. *You guys* must think I’m a quitter, to be scared out of a little old church show (Ade 130).

In (13), *guys* functions to designate, or further identify, the other two men in the scene as the plural referents of *you*, while Artie engages them in the shared challenge. As I show

in Chapter 6, the original masculine appositive function of *you guys* in Ade's example no longer applies to many of *you guys*'s addressive contexts today.

Based on the COHA, Table 9 lists the ten most frequent [PRO (AP) AppNP] uses with *you* and a plural common noun prior to the appearance of *you guys* in Ade's novella, both with and without the optional AP:

Example (without AP)	Token frequency	Example (with AP)	Token frequency
<i>you men</i>	150	<i>you young folks</i>	12
<i>you fellows</i>	89	<i>you young men</i>	10
<i>you boys</i>	76	<i>you young ladies</i>	10
<i>you women</i>	67	<i>you young fellows</i>	5
<i>you gentlemen</i>	60	<i>you young girls</i>	4
<i>you girls</i>	46	<i>you young women</i>	4
<i>you ladies</i>	45	<i>you American ladies</i>	4
<i>you children</i>	27	<i>you lazy rascals</i>	3
<i>you Americans</i>	24	<i>you little rascals</i>	3
<i>you dogs</i>	24	<i>you young rascals</i>	3

Table 9: The ten most frequent examples of [*you* (AP) AppNP] prior to *you guys*'s attestation, with and without an AP, 1810-1880 (COHA)

You men and *you fellows* were the most frequent *you*-appositives before *you guys* emerged in 1896, but other variants were prevalent in the 1800s. The initial usage of *you guys* represented the increasing productivity of the [PRO (AP) AppNP] construction as it accepted newer types of common plural nouns.

The uses with APs on the right side of Table 9 illustrate the three distinct slots of the appositive construction, demonstrating the occasional non-adjacency of *you* and the following NP. While *you young folks* was the most frequent example of such a use in the 1800s, *folks* entered the construction much less frequently when it lacked an intervening AP, hence *you folks*'s exclusion in the leftmost column. *You men*, by contrast, was relatively high frequency in the 19th century both with and without an intervening AP.

You fellows was likewise high frequency among *you*-appositives, especially when lacking an intervening AP.

Given the similar lexical basis of *you men*, *you fellows*, and *you guys* in the [PRO (AP) AppNP] construction and their overlapping masculine meanings, the former two variants were in principle more viable candidates for grammaticalization prior to turn of the century, being much higher in frequency. The question naturally arises, then: Why did it turn out to be *you guys* that was gradually reshaped into a semantically general 2p pronominal unit, but not *you men* or *you fellows*? In the course of the present study, I address this question by exposing the distinctive usage pattern of *you guys* in the 20th century, along with its unique structural and semantic-pragmatic changes within the [PRO (AP) AppNP] construction.

4. The role of frequency in the development of *you guys*

If repetition, measured by token frequency, can be shown to influence the development of pronominal forms, then we have evidence against the traditional view that grammar is autonomous from usage (Section 1.2); conversely, we have evidence in support of a usage-based theory of grammar. In the particular case of *you guys*, where we encounter frequency gains and their predictable effects on grammaticalizing forms, we can be more confident in the claim that it has grammaticalized, having moved beyond its primary stage as a [PRO (AP) AppNP]. The central question of this chapter is whether *you guys* has gained significantly in frequency since its appearance in American English. Is *you guys*'s frequency positively correlated with its decade of usage, as would be expected if grammaticalization were taking place?

Though Ade's usage is earlier by 15 years, in the COHA *you guys* first appears in 1911 in Edna Ferber's novel *Dawn O'Hara*. In the following scene, Ferber's character Blackie, who has been hospitalized, comments on the lack of enthusiasm that his two male friends, Norberg and Deming, have inspired during their visit to the hospital:

(14) Guess *you guys* ain't got th' stimulin' effect that a bunch of live wires ought to have (COHA, Dawn O'Hara, 1911).

Ferber's example is one of six *you guys* tokens in the 1910s period of the COHA, each involving a different author. In other words, *you guys* does not emerge as an idiosyncratic feature of one individual's writing along the lines of a nonce form, but rather as a broader linguistic phenomenon in the corpus.

Figure 1 details changes in the frequency of *you guys* from its earliest wave of usage in the works of Ferber and others through the present day, based on a search for the

you guys string in the COHA (i.e. omitting the *guys*-vocative and tokens with intervening adjectives or punctuation). Because this corpus includes only written texts for each period, it enables us to subtract the effect of *you guys*'s higher frequency in more interactive spoken registers (Section 4.3), which might otherwise bias its frequency toward recent periods.

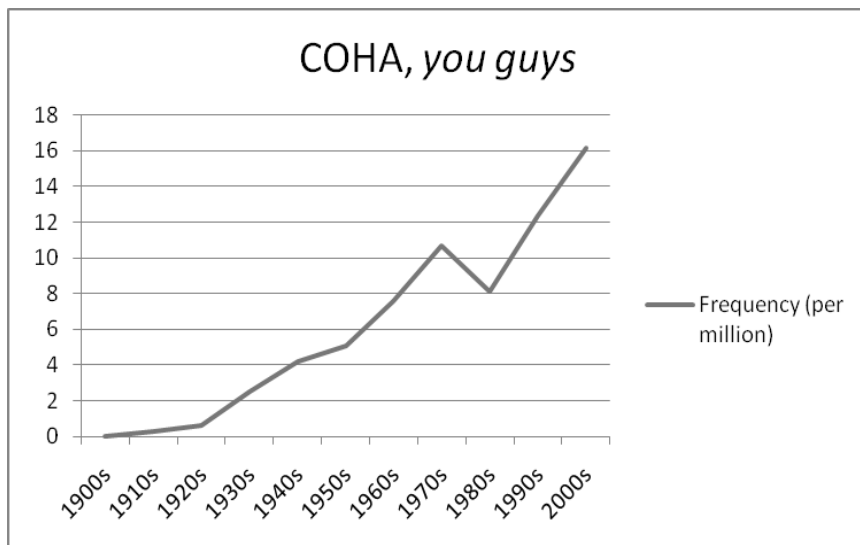


Figure 1: The normalized frequency of *you guys*, by period, COHA (Token frequencies: 1900s: 0; 1910s: 6; 1920s: 16; 1930s: 62; 1940s: 102; 1950s: 125; 1960s: 182; 1970s: 255; 1980s: 205; 1990s: 345; 2000s: 478; All: 1,776)

During the decade of Ferber's example from (14), or the 1910s, the normalized frequency of *you guys* was .26 per million words. The string is 62 times more frequent today, however, having increased sharply in its rate of occurrence to 16 per million words (2000s). A Kendall's *tau* test confirms that the positive correlation between frequency and decade of usage in Figure 1 is strong and extremely significant at $p < .05$ ($p = .0001$, $\tau = .97$, two-tailed). The likelihood of such a relationship in the corpus being due to chance is exceedingly low at .01%.

The Kendall's *tau* test is designed for the verification of dependencies between two variables (here frequency and decade) in abnormally distributed data sets such as that

of Figure 1, in which the frequency of *you guys* is severely skewed toward recent decades. For this sort of computation, Kendall's *tau* yields more conservative and reliable results on average compared to alternative tests that assume normality, such as Pearson's *r* (Gries to appear: 16, personal communication).

4.1 A comparison of *you*-appositives

To strengthen my claim that *you guys* has grammaticalized via structural and semantic changes in the [PRO (AP) AppNP] construction, it is additionally necessary to compare the usage of *you guys* to other instances of this construction. A view across instances of the appositive construction can demonstrate the uniqueness of *you guys*'s diachronic frequency distribution and related structural and semantic changes. Such a line of inquiry requires three main assessments: (1) the diachronic frequency distributions of comparable instances of the [PRO (AP) AppNP] construction, (2) any changes in the constituency of these instances, and (3) any changes in the meanings of these instances. In Chapter 4, I address the first issue, *you guys*'s frequency of usage compared to the other *you*-appositives *you men* and *you fellows*, including the related structure *you all*. Chapter 5 covers the second topic, the selective chunking of *you guys* with respect to these other variants, and Chapter 6, finally, looks into the semantic changes of *you guys* relative to *you men* and *you fellows*, which have similar semantic origins.

As Table 9 has indicated, before *you guys* entered American English, there were a number of [PRO (AP) AppNP] uses with masculine semantics, including *you men*, *you fellows*, *you boys*, and *you gentlemen*, in descending order of frequency. The following examples are from 1880s portion of the COHA:

(15) “Pooh!” said Sybil; “*you men* are all just alike...” (COHA, Democracy American).

- (16) “Look here, if *you fellows* keep interrupting, I won’t sit down for half an hour...”
(COHA, Jack Jill).
- (17) “*You boys* don’t seem to have his appetite for liquor. You are a member of Congress, and Elk was one of the bravest ginerals [sic] in the war...” (COHA, Tales Chesapeake).
- (18) “It is all very well for *you gentlemen* to measure General Washington according to your own private twelve-inch carpenter’s rule. But what will you say to us New Englanders who never were country gentlemen at all...” (COHA, Democracy American).

To narrow the scope of my comparison between *you guys* and this set of *you*-appositives, I center on the two most frequent variants, *you men* and *you fellows* (Table 9). In additional comparisons, I discuss *you all* and *you-p* to underscore the particular trend in usage that *you guys* has followed throughout the 20th century.

4.1.1 The diachronic frequency of *you men* and *you fellows*

As Figures 2 and 3 reveal, one principal difference between the diachronic frequency of *you guys*, *you men*, and *you fellows* is that the latter two variants never made significant headway in usage, instead experiencing transitory popularity early in the 20th century. Figure 2 details the frequency of *you men* in the COHA:

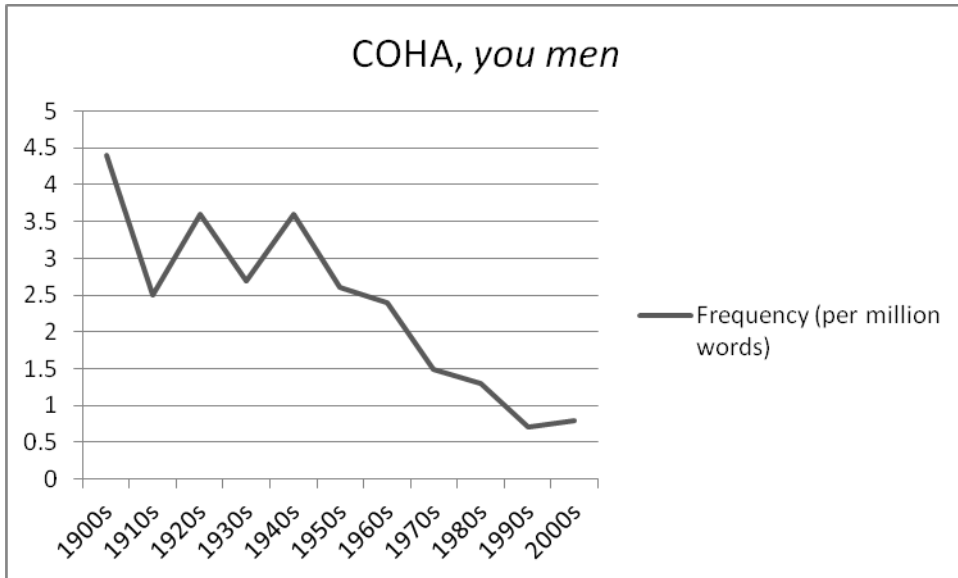


Figure 2: The normalized frequency of *you men*, by period, COHA (Token frequencies: 1910s: 56; 1920s: 91; 1930s: 67; 1940s: 87; 1950s: 83; 1960s: 57; 1970s: 36; 1980s: 33; 1990s: 19; 2000s: 23; All: 552)

Based on a Kendall's *tau* test, the interdependence of frequency and decade of usage in

Figure 2 is strong and extremely significant at $p < .05$ ($p = .001$, $\tau = -.8$, two-tailed), yet contrary to *you guys*, this relationship is negative. A juxtaposition of Figure 1 and Figure 2 reveals that in the 1910s *you men* was nearly 10 times more frequent than *you guys*, but by the end of the 20th century this trend had been reversed, with *you guys* being 15 times more frequent than *you men*.

A similar tendency holds for *you fellows* in the COHA, as Figure 3 illustrates:

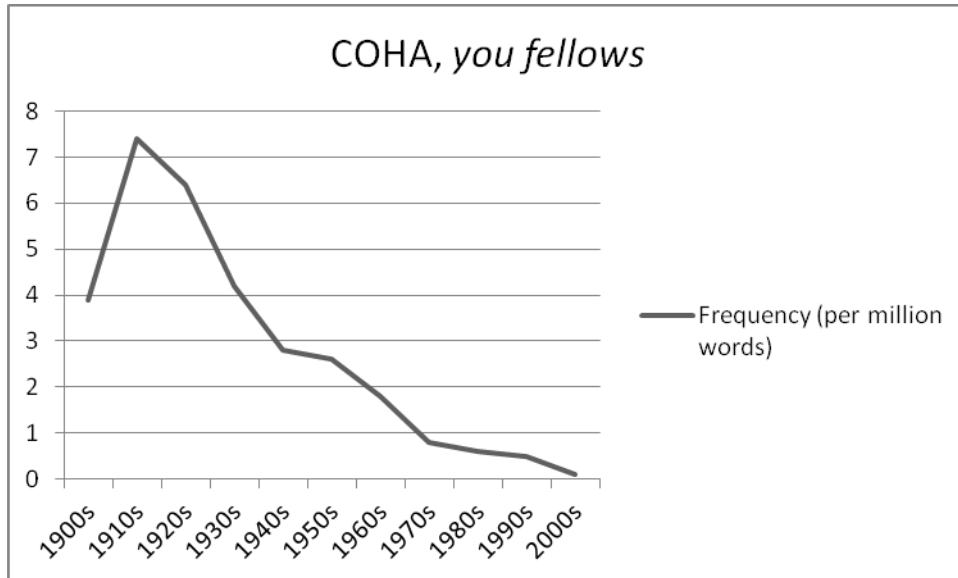


Figure 3: The normalized frequency of *you fellows*, by period, COHA (Token frequencies: 1910s: 167; 1920s: 165; 1930s: 102; 1940s: 67; 1950s: 63; 1960s: 43; 1970s: 19; 1980s: 14; 1990s: 14; 2000s: 4; All: 658)

A Kendall's *tau* test establishes an even stronger and more statistically significant negative correlation between the frequency and decade of *you fellows* usage ($p=.0001$, $\tau=-.9$, two-tailed), which underscores *you guys*'s distinctly positive trend in usage in the 20th century. Whereas in the 1910s *you fellows* was almost 29 times more frequent than *you guys*, by the 2000s, *you guys* had become 162 times more frequent than *you fellows*.

Given the relationship between usage and language change clarified in Section 1.2, the sharply decreasing frequency of *you men* and *you fellows* has corresponded to their failure to grammaticalize as pronominal units. Meanwhile, *you guys*'s precipitous gain in frequency has promoted the processes of grammaticalization—and the form has become even more frequent because of those processes—as Chapters 5 and 6 demonstrate.

4.1.2 The frequency of *you all*

You all (including its related orthographic variants *you-all*, *ya'll*, and *y'all*) has followed a more peculiar trajectory in American English with respect to *you guys*, *you men*, and *you fellows*. First, the *you all* appositive construction differs slightly in its original structure, incorporating a quantifying PRO in the NP slot that was unmarked for number and already semantically general, or gender neutral: [PRO (AP) AppPRO].

Words meaning ‘all’ are also Second, as Figure 4 illustrates, the frequency of *you all* has always been higher than *you guys*, *you men*, and *you fellows*, though *you all*’s frequency has remained much flatter over time:

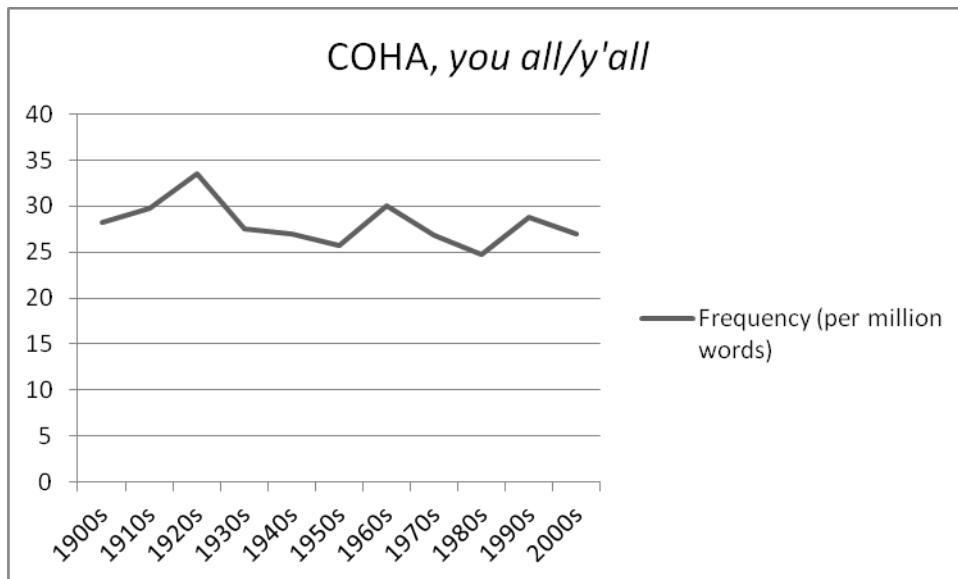


Figure 4: The normalized frequency of *you all/y'all*, including *you-all* and *ya'll*, by period, COHA, 20th century (Token frequencies: 1910s: 674; 1920s: 858; 1930s: 674; 1940s: 657; 1950s: 631; 1960s: 719; 1970s: 641; 1980s: 623; 1990s: 806; 2000s: 831; All: 7,114)

Thus the dependency between the frequency and decade of *you all* usage is weak and does not achieve statistical significance at $p < .05$ ($p = .16$, $\tau = -.3$, two-tailed). The higher frequency of *you all* with respect to other appositive uses and its simultaneous flatness in frequency over time indicate an earlier frequency gain, sometime prior to the 1900s. The

1800s portion of COHA offers little help in pinpointing exactly when this increase occurred:

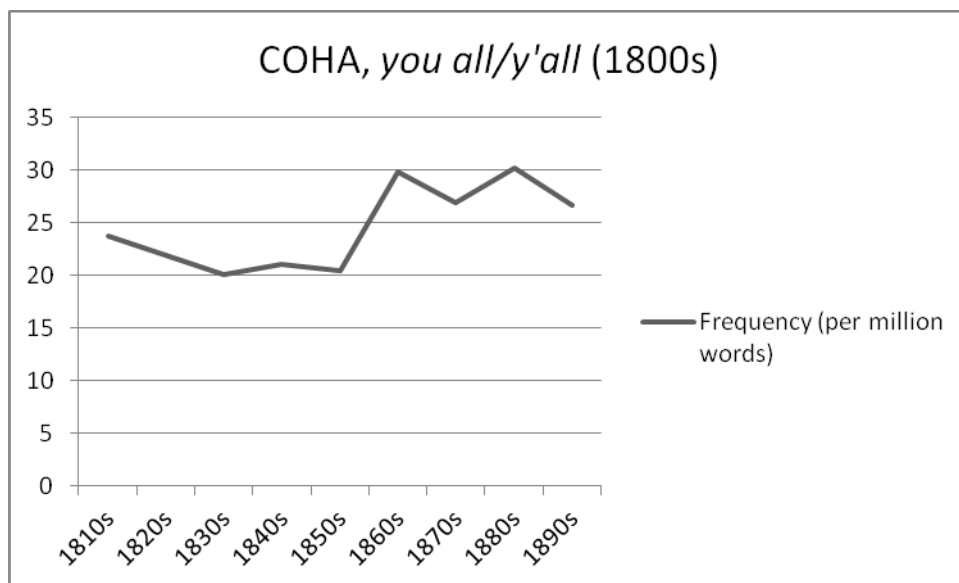


Figure 5: The normalized frequency of *you all/y'all*, including *you-all* and *ya'll*, by period, COHA, 19th century (Token frequencies: 1810s: 28; 1820s: 152; 1830s: 278; 1840s: 340; 1850s: 337; 1860s: 509; 1870s: 504; 1880s: 615; 1890s: 650; 1900s: 736; All: 4,149)

While on the surface a frequency gain for *you all* appears to have occurred in the 1850s, the frequencies for all prior periods remain quite high relative to other instances of the appositive construction. The relationship between frequency and decade of usage, moreover, is somewhat weak and fails to reach significance at $p < .05$ ($p = .1$, $\tau = .4$, two-tailed). It is probable, then, that *you all*'s initial ascent in usage occurred during or before the 18th century, since then the form remaining relatively stable in usage. Because the COHA reaches back only as far as the 19th century, however, we cannot be certain of when this increase took place. What we can safely conclude is that the usage of *you all* in the 20th century departs from the “hockey stick-like” pattern of a form whose grammaticalization is becoming progressively more rapid; this pattern better characterizes the frequency distribution of *you guys* in Figure 1.

4.1.3 Summary and discussion

In light of the foregoing frequency data, it is clear why today *you all* and *you guys* far outstrip other instances of the appositive construction in usage:

Use	Token Frequency
<i>you all/y'all</i>	831
<i>you guys</i>	478
<i>you boys</i>	91
<i>you kids</i>	41
<i>you girls</i>	32
<i>you men</i>	23
<i>you folks</i>	21
<i>you Americans</i>	16
<i>you gentlemen</i>	11
<i>you fools</i>	9

Table 10: Today's top ten most frequent *you*-appositive uses, COHA (2000s)

The standing of *you all* in the table reflects its history as the highest frequency instance of the appositive construction throughout the 19th and 20th centuries. As for *you guys*, it has instead been surging in frequency since the early part of the 20th century, effectually climbing its way toward the top of Table 10.

The uniquely positive trend in the frequency of *you guys* in the 20th century compared to other *you*-appositives suggests that its development is exceptional. While the most frequent *you*-appositives at the start of this century later decline or flatten in frequency, *you guys* spikes dramatically in usage and continues to do so into the present day. Insofar as the processes of grammaticalization both depend on and propel usage, the frequency distributions of *you guys* and other *you*-appositives are crucial in explaining why grammaticalization has occurred “selectively,” operating specifically on *you guys* but not on other constructions such as *you men* and *you fellows*. As for *you all*, though it may be the case that it continues to grammaticalize in PdAE in the face of its higher

frequency and apparent structural changes (Section 4.1.2), its lack of a frequency gain in the 20th century implies that it is developing less rapidly than *you guys* PdAE.

4.2 A frequency comparison of *you guys* and *you-p*

In addition to a comparison between the diachronic frequency of *you guys* and other *you*-appositives, it is important to relate *you guys* to *you-p* in order to rule out the possibility that *you guys*'s frequency gains reflect a more general increase in American English address.

Another benefit of examining *you-p* usage is that it permits an assessment of dialect surveys such as the Harvard Survey of North American Dialects (Vaux 1999/2005), which rank participant preferences for 2p variants. To what extent do reported preferences in the Harvard survey coincide with actual 2p usage as represented in the COHA and COCA? Can this sort of survey be taken as a reliable indicator of American English address?

In large-scale corpora such as the COHA and COCA, a frequency comparison between *you-p* and other 2p variants can nevertheless be thorny due to *you*'s structural ambiguity across its number-marking functions. Unsorted *you* tokens in these corpora number in the millions, rendering it methodological unreasonable to parcel out all plural tokens.

To circumvent this obstacle, ratio-based tallies have been performed on a sample of *you* tokens from the corpora (Section 4.2.1). While the COHA allows for a diachronic comparison between *you-p* and other 2p variants to discount the possibility that 2p usage has increased overall (i.e. irrespective of *you guys*'s particular increase), the COCA

allows for a more legitimate evaluation of the Harvard survey results, since participants in the survey reported on what they would say, not write.

4.2.1 The frequency of *you-p* in the COHA

The ratio-based count of *you-p* in the COHA involved three primary steps. First, using the corpus’s randomizing “Sample” function, I collected 100 *you* tokens per decade (20th century), yielding 1000 tokens overall. Each set of 100 decade-specific *you* tokens represented 25 per genre—fiction, magazines, newspapers, and non-fiction—to mitigate against genre-bias.

Second, I checked the 1000 expanded contexts associated with these tokens and sorted *you* number-marking functions into one of three categories: singular referent (*you-s*), plural referents (*you-p*), or indefinite/other (*you-i/o*).² Examples of each of these uses from the sample follow:

- (19) Singular: “Why, Emil! I told *you* to stay in the store and not come out”
(COHA, *O Pioneers!*, 1913).
- (20) Plural: “She says she used to play with *you* when *you* were children”
(COHA, *Age of Innocence*, 1920).
- (21) Indefinite: “If *you* pry away some of the wall to spy on them [termites], *you* get the fiasco I was just rewarded with” (COHA, *Raid on Termites*, 1932).

In indefinite uses like (21), *you* conveys a hypothetical meaning instead of functioning deictically to spotlight a particular set of participants relative to the speaker; indefinite

² I am employing Lyons’ (1999) notion of indefiniteness here, detailed in Section 6.2. Here and elsewhere, the term ‘generic’ has been avoided to bypass confusion with *you guys* uses having generic gender and yet specific, definite reference (Sections 6.3-6.5).

you is therefore ambiguous between a singular and plural meaning, or semantically neutral with respect to number (see Section 6.2 for further discussion of indefiniteness). Tokens from the sample constituting “Other” *you* uses have been grouped with the indefinites due to their similar non-referential quality, including sequences such as *you see*, *you know*, and *thank you*, where *you* does not function to profile a specific set of interlocutors in the discourse.

Third, for each relevant time period, I counted the tokens in the plural category to determine the percentage (ratio) of *you*-p tokens relative to others. 95% confidence intervals were then computed to determine the range in which the true proportion of *you*-p uses likely falls in the COHA, given the estimated sample proportions and sizes for each period. Lastly I multiplied the sample proportions (percentages) by the total number of unsorted *you* tokens, period by period. This process yielded a numerical estimate of *you*-p tokens per period, which was subsequently normalized into a frequency per million words to allow for valid comparisons across 2p variants and corpora.

Results of the *you* number sort are given in Table 11:

Period	Unsampled <i>you</i> tokens*	Sampled <i>you</i> tokens	<i>you</i> -s	<i>you</i> -p	<i>you</i> -i/o	% <i>you</i> -p (sample)	CI**	Total <i>you</i> -p tokens (est.)	Normal-ization factor	Frequency per million words, <i>you</i> -p (est.)
1910	182,490	100	55	6	39	6	4.7	10,949	22.7	482
1920	182,286	100	68	5	27	5	4.3	9,114	25.6	356
1930	172,206	100	52	6	43	6	4.7	10,332	24.4	423
1940	169,999	100	62	8	30	8	5.3	13,600	24.1	564
1950	184,208	100	47	6	47	6	4.7	11,052	24.4	453
1960	174,152	100	43	7	50	7	5	12,191	23.9	510
1970	177,234	100	39	8	53	8	5.3	14,179	23.8	596
1980	179,555	99	33	1	64	1	2	1,796	25.2	71
1990	211,167	100	38	2	60	2	2.7	4,223	27.9	151
2000	219,008	100	37	3	60	3	3.3	6,570	29.5	223
Overall	1,852,305	999	474	52	473	5.2	1.4	94,007	N/A	AVG 383

Table 11: Portion and frequency estimates for *you*-p relative to other *you* uses, COHA, 20th century (*=excluding *you all* and *you guys* tokens, **=95% confidence intervals for % *you*-p, based on sample proportions and sizes, corrected for actual population sizes, i.e. unsampled *you* tokens)

In each respective decade and in the sample overall, the overwhelming majority of *you* tokens are either singular or indefinite/other. A much smaller portion of *you* tokens are plural, ranging from 1 to 8 percent depending on the decade, with an overall sample proportion of 52/999, or 5.2 percent. With a confidence interval for this proportion of 1.4, there is a 95% probability that the segment of *you*-p uses relative to other *you* instances in the COHA falls somewhere between 3.8 and 6.6 (20th century).

Even at the lower end of this 3.8-6.6 range, however, the frequency of *you*-p in the COHA far exceeds that of *you all*/*y'all* and *you guys*. Refer to Table 12:

	Token frequency		
	<i>you</i> -p (est.*)	<i>you all</i> / <i>y'all</i>	<i>you guys</i>
1910-2009	70,388	7,114	1,776

Table 12: The frequency of *you*-p vs. *you all* and *you guys*, COHA (*=based on lower limit of confidence interval for overall percent *you*-p in sample, see Table 11)

With a maximally conservative token frequency estimate of 70,388 in the COHA, *you-p* is used nearly 10 times more often than *you all/y'all* and 40 times more often than *you guys*. Thus while *you all/y'all* and *you guys* are higher in frequency than the other *you-* appositives in the corpus (Table 10), they hardly approach the frequency of *you-p*. This fact puts into perspective the embryonic nature of *you guys*'s grammatical development, lacking the long-standing history in the language of *you-p* (Section 1.1).

It is clear, nevertheless, that while *you guys* has gained significantly in frequency over the 20th century in the COHA (Figure 1), *you-p* has not. In this regard Figure 6 displays diachronic changes in the estimated normalized frequency of *you-p* from the COHA sample, based on the final column of Table 11:

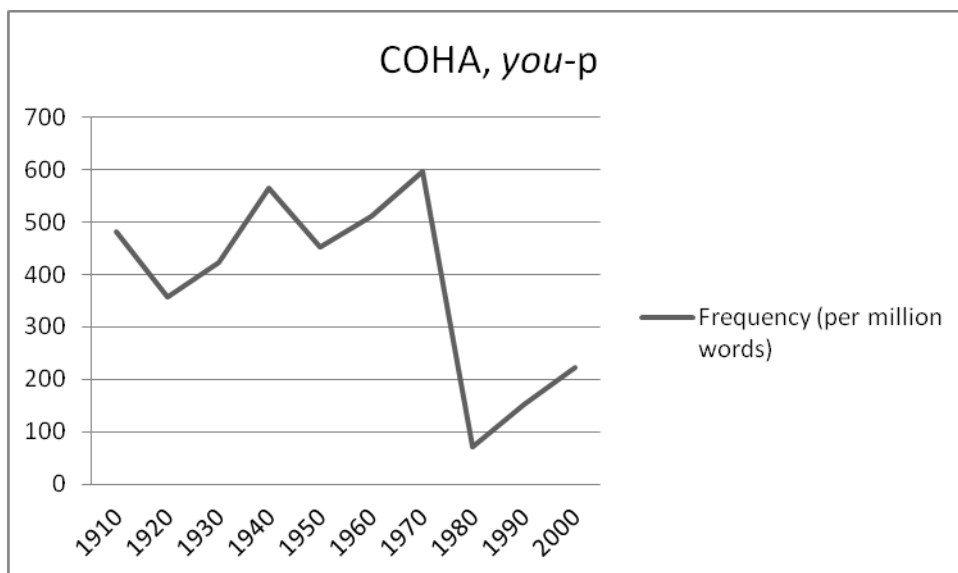


Figure 6: Estimated frequency of *you-p*, by period, COHA (based on Table 11)

It appears superficially that *you-p* has declined in frequency over roughly the same period in which *you guys* has made its most dramatic increase in usage, the latter half of the 20th century. Were this the case, we would have evidence that *you guys* (among other 2p variants) is cutting into the usage of *you-p*. The relationship between *you-p*'s frequency

and decade of usage, however, remains statistically insignificant at $p < .05$ ($p = .53$, $\tau = -.2$, two-tailed). Thus it is uncertain whether the negative relationship in the figure reveals a meaningful trend in *you-p* usage.

What we can conclude from Figure 6 with greater certainty is that while *you guys* has expanded in usage in the 20th century, *you-p* has not made similar frequency gains; like *you all*, the frequency of *you-p* has instead been relatively stable over time. This contrast highlights the uniquely positive trend in usage that *you guys* has entertained in PdAE, discounting the hypothesis that the form has followed such a trend by virtue of a more general increase in 2p address throughout the 20th century.

The following section confirms that the increasing usage of *you guys* encountered in the COHA is not particular to writing, or genre-specific, and that the form is in fact more common in speech.

4.3 *You guys*'s higher frequency in speech: the COCA data

Similar to the written data for recent periods (COHA, 1990s-2000s), a sizeable increase in *you guys*'s frequency can be found in the spoken section of the COCA from 1990 forward. Refer to Figure 7:

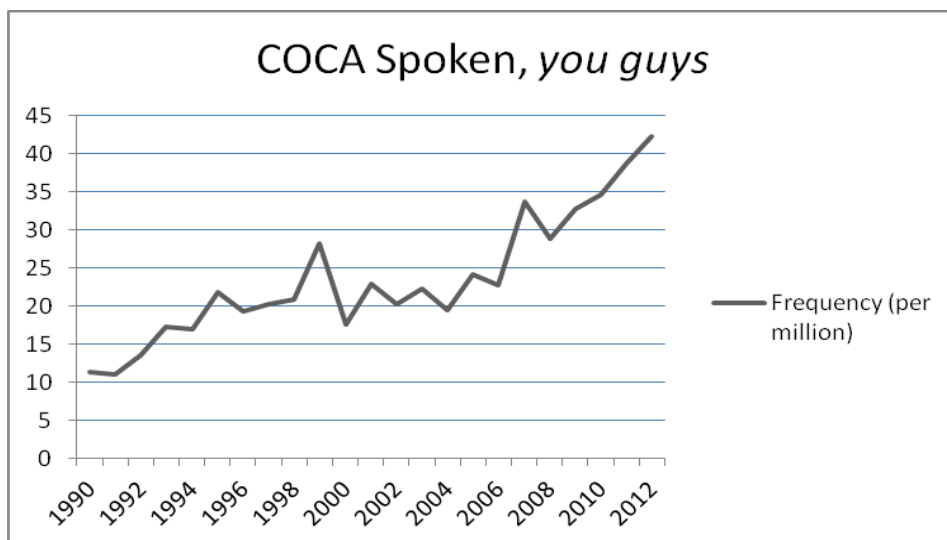


Figure 7: The normalized frequency of *you guys*, by year, COCA Spoken (Token frequencies: 1990: 232; 1991: 227; 1992: 282; 1993: 361; 1994: 358; 1995: 456; 1996: 394; 1997: 415; 1998: 431; 1999: 592; 2000: 366; 2001: 460; 2002: 414; 2003: 463; 2004: 405; 2005: 503; 2006: 473; 2007: 685; 2008: 578; 2009: 658; 2010: 688; 2011: 797; 2012: 481; All: 7,220)

In speech, the usage of *you guys* more than doubles from 1990 to 2009, a frequency gain that is even more striking than that in writing. From 1990 to 2012, *you guys*'s rate of occurrence in the spoken section of the COCA increases by 252%, from 11 to 42 instances per million words. Along with the written data, the positive correlation between *you guys*'s frequency and period of usage in Figure 7 is strong and extremely statistically significant at $p < .05$ ($p = .0001$, $\tau = .8$, two-tailed). This fact corroborates the proposal of Sections 4.1 and 4.2 that *you guys* has been uniquely positioned to grammaticalize in PdAE, in contrast to other 2p variants whose usage has slumped or stalled over time.

You guys is also more frequent in the COCA spoken section overall than in comparable sections of the COHA (1990s and 2000s). In the most recent decades of the COHA the normalized frequency of *you guys* is 14 on average, whereas in the COCA its mean rate of occurrence is 24 per million words, making it twice as common in speech than in writing. This difference can be reconciled by the facts that (a) second-person

address occurs more often in spoken language, as the communicative mode includes more interaction among participants, and (b) written language can be more conservative than spoken language, with editors enforcing revisions (e.g. the substitution of *you guys* for another form). *You guys*'s higher frequency in speech suggests that its grammaticalization is unfolding more quickly in that communicative mode, while its mutual frequency gain in writing and speech implies that its grammaticalization is not ultimately restricted by modality.

In the next section, I compare the frequency of *you guys* and other 2p variants in contemporary spoken American English and discuss the challenge that the corpus data pose to existing surveys of 2p address.

4.4 Reconciling corpus frequencies and survey rankings of 2p addressives

The positive trend in *you guys* usage in the written and spoken data coincides with a broad public perception in the United States of its acceptability in 2p address. In fact *you guys* is the most preferred 2p variant on the Harvard Survey of North American Dialects (Vaux 1999/2005). Question (50) of this survey asks, "What word(s) do you use to address a group of two or more people?" Participants then select options from a list, with the ability to choose multiple options if they deem necessary.

In response to this question, *you guys* is strongly favored over all other 2p variants:

2p variant	Percentage of participants who chose the variant for Question (50)
<i>you guys</i>	43
<i>you</i>	25
<i>y'all</i>	14
<i>you all</i>	13
<i>other</i>	5
<i>yous(e)</i>	1
<i>yins</i>	<1
<i>yuns</i>	<1
<i>you lot</i>	<1

Table 13: Harvard Survey of North American Dialects (n=10,764)

According to Table 13, 18% more participants choose *you guys* over *you-p* in response to Question (50) of the survey, and this disparity is still greater for the other 2p variants below *you* in the table.

It is evident, moreover, that the participants who selected *you guys* for Question (50) are wide-ranging in their areal distribution. In this respect Figure 8 shows the regional distribution of 2p preferences from the survey:

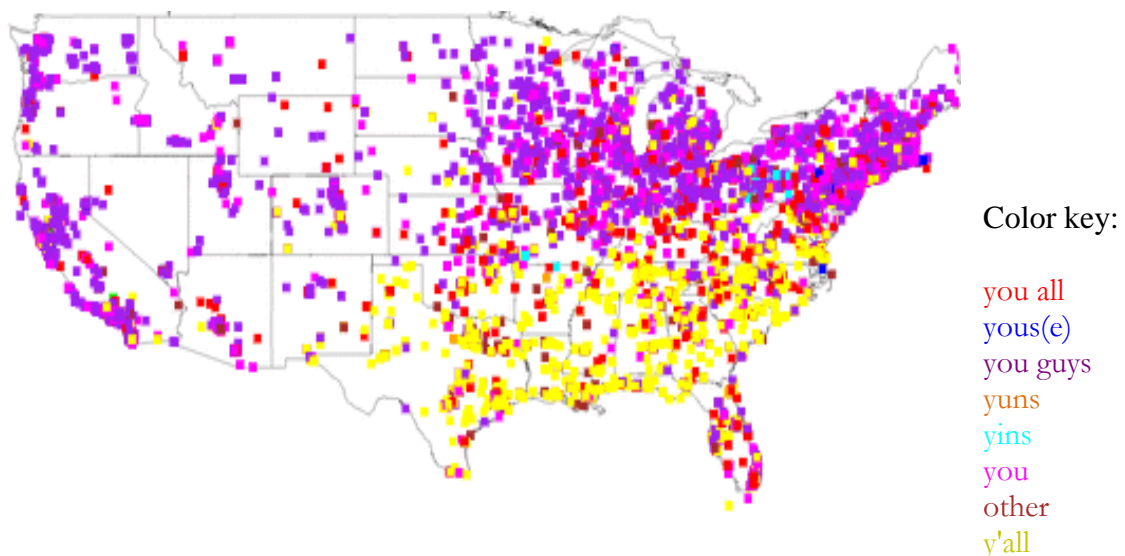


Figure 8: Regional distribution of 2p preferences (Harvard Survey of North American Dialects)

Together with *you-p* and *you all*, *you guys* is much more broadly distributed than the other variants in the figure, being associated with the West, Midland, Northern, Mid-Atlantic, and New England regions, in addition to select areas in Florida and the South. Thus it is feasible to classify *you guys* as an ‘American’ English addressive overall, as opposed to a regionalism.

Along these lines, Maynor 2000 reports that *you guys* is moving into the South based on a survey of university students in Mississippi, Alabama, North Carolina, and South Carolina; this finding sits well with the results of the Harvard survey, which reveal *you guys* preferences in some Southern areas. Through their telephone focus poll, Tillery et al. 2000, meanwhile, demonstrate that *y’all* is moving out of the South. At first glance, its selection in the Harvard survey remains concentrated in the South (see yellow); on closer inspection, however, one finds that *y’all* is chosen by participants in the West, North, Midland, Mid-Atlantic, and New England regions, albeit more spottily than *you guys*. The more limited areal distribution of *y’all* outside of the South is perhaps due in part to hesitance among participants to report using the form, given the persistence of its social stigmatization in the United States. In Section 5.1.2, I demonstrate that *y’all* has in fact been increasing significantly in frequency since the 1980s, similar to *you guys*, but only when counted separately from *you all* (contra Figure 5). There is a growing body of evidence, for this reason, that *y’all* too is becoming nationalized.

In addition to the trans-regional status of *you*, *you guys*, *you all*, and to some degree *y’all*, Figure 8 shows that these variants overlap considerably in their area of selection. In locales in which this occurs, the variation among the forms must be attributed to factors other than geography. On the one hand, this geographically-specific

variation may be due to the diverse compositions of local populations (i.e. dialect mixing), given the rise of interstate migration in the United States noted in demographic research (Rosenbloom and Sundstrom 2003). A common effect of this sort of demographic change is dialect leveling (Kerswill 2003), or the diminishment of regional distinctiveness in speech. On the other hand, the variation in 2p forms may be due to style-shifting (changes in register) or to the immediate structure of the discourse. This scenario further underscores the necessity to identify the discourse constraints on 2p usage, as done later for *you guys* in Chapter 6.

4.4.1 Corpus frequencies of contemporary 2p addressives

Due to the occasional inaccuracy of self-perceptions in linguistic surveys, it becomes necessary to evaluate Table 13 in relation to the contemporary corpus data. To this end, Table 14 shows the frequency of 2p variants in the COCA, representing those from the Harvard survey in addition to two others discussed in the literature, *you people* and *you folks*, recorded as ‘Other’ in the survey.

COCA Spoken (1990-2012)			COCA Overall (1990-2012)		
2p variant	Frequency (per million)	Token Frequency	2p variant	Frequency (per million)	Token Frequency
<i>you-p*</i>	1,536	147,457	<i>you-p*</i>	308	142,776
<i>you guys</i>	75.6	7,220	<i>you all & y'all</i>	25.6	11,879
<i>you all & y'all</i>	65.8	6,282	<i>you guys</i>	23.3	10,829
<i>you all</i>	62	5,923	<i>you all</i>	22.4	10,391
<i>you people</i>	7.42	709	<i>y'all</i>	3.2	1,488
<i>y'all</i>	3.8	359	<i>you people</i>	3.1	1,431
<i>you folks</i>	3.5	334	<i>you folks</i>	1.2	567
<i>yous(e)</i>	0.5	50	<i>yous(e)</i>	0.4	166
<i>you lot</i>	0.02	2	<i>you lot</i>	0.1	30
<i>yins/yinz</i>	0	0	<i>yins/yinz</i>	0.002	1
<i>yuns/yunz</i>	0	0	<i>yuns/yunz</i>	0	0

Table 14: The frequency of contemporary 2p variants (COCA) (*=row numbers derived from ratio-based tallying method described in Section 4.2.1, adapted to the COCA)

A number of conclusions can be drawn from Table 14. For one, the Harvard survey is an accurate indicator of American English usage to the extent that the highest frequency group of 2p variants includes *you-p*, *you guys*, *you all*, *y'all*, and other (e.g. *you people*).³ Secondly, the survey inaccurately predicts the individual frequency rankings of these variants, inflating the ranking of *you guys*, for instance. In both the spoken portion of the COCA and in the corpus overall, *you-p* in fact remains the most frequent form by a considerable margin, especially in speech.

Modality does not impact the frequency ranking of each variant uniformly, however. For example, *you guys* outnumber *you all/y'all* in speech, while in the COCA

³ One caveat to Table 14 is that the lower frequency of *y'all* with respect to *you people* in the leftmost column (spoken) may in part be attributable to transcription practices underlying the corpus; that is, many *y'all* tokens could have been conceivably entered as *you all*. A similar case in which transcription practices most likely obfuscate frequency in the COCA involves “I don’t know,” which is often reduced in speech (Bybee and Scheibman 1999). While the string *I dunno* turns up just 11 times in the corpus overall, *I don’t know* records 85,925 hits. Given what we know about this use (ibid.), it is highly probable that many of *I dunno* tokens wind up being encoded as *I don’t know* in COCA transcripts.

overall (including written texts) this ranking is reversed, mirroring the ranking from the COHA (Chapter 4). Again, this discrepancy illustrates the need to examine the situational attributes of 2p address more carefully.

The sharp decrease in frequency below *you all* in Table 14 implies that the dominant 2p variants of American English—or those with the strongest lexical representations across speakers represented in the COCA—are *you-p*, *you guys*, and *you all/y'all*. The co-existence of *you-p*, *you guys*, and *you all/y'all* at relatively high frequencies in 2p address exemplifies the concept of “layering” in grammaticalization (Hopper 1991), by which different forms serve largely overlapping functions while a particular area of the grammar undergoes change. The layering of 2p addressives in American English suggests that the pronominal system is undergoing yet another wave of historical changes (recall Section 1.1), shown especially by the frequency distributions of *you guys* and *y'all*.

As Hopper 1991 argues, layering is one of five phenomena in language change that can be used to gauge whether grammaticalization is taking place, its other hallmarks being decategorialization, divergence, persistence, and specialization. Later I reveal how these other indicators of grammaticalization are likewise applicable to *you guys* by relating them to changes in its constituency and meaning.

4.4.2 Summary and discussion

Chapter 4 has demonstrated the substantial increase in *you guys*'s frequency throughout the 20th century (Figure 1) and its continued expansion in usage today, both in writing and speech (Figures 1 and 7). Thus a key criteria of grammaticalization has been met in the corpora: significant frequency gains.

The distinctive diachronic usage of *you guys* with respect to other 2p variants illustrates the exceptional nature of its development throughout the 20th century. While *you guys* nowhere approaches *you-p* in usage as erroneously predicted by the Harvard Survey of North American Dialects, *you guys* is nevertheless one of the top three most frequent 2p addressives in American English today (Table 14); in speech, only *you-p*, in fact, is used more often than *you guys*. The relatively high frequency of *you guys* among 2p addressives coincides with its broad areal distribution in the United States. Given this distribution, it is unlikely that *you guys*'s frequency is biased by highly concentrated usage in one particular region to the exclusion of others, such that *you guys* would be better characterized as a regionalism. The evidence thus far uncovered is well-catered to the hypothesis that *you guys* has been grammaticalizing in the 20th century and continues to grammaticalize across large stretches of the United States.

While it has been shown that the self-perceptions embodied in the Harvard survey are in many ways inaccurate predictors of 2p frequency rankings in actual usage, it is reasonable to assume that survey respondents are more accurate at noting the mere presence of a 2p variant in their speech. This self-perception can be paraphrased as 'Yes, I use this word/phrase' as opposed to 'Yes, I use this word/phrase the most/the least/more than that one' or 'Yes, I use this word/phrase instead of that one'. The latter two self-perceptions are, in my view, more problematic in predicting usage by requiring ranked assessments of multiple forms. Though the Harvard survey is less effective at foretelling the frequency rankings of 2p variants in usage, it is more effective, then, at capturing the simple existence of a variant in a speech community. For this reason the survey can be

trusted in a limited sense as an indicator of *you guys*'s relative ubiquity in the United States.

A juxtaposition of Figure 8 and Table 14 clarifies that the 2p variants with ties to the most aurally specific regions in the Harvard survey—*yous(e)*, *yuns*, and *yins/yinz*—are in fact lowest in frequency in the COCA. The higher frequency of *you-p*, *you guys*, and *you all/y'all* in the corpus corresponds conversely to their broader areal distribution in the survey. This relationship further stresses the fact that *you guys* is an addressive of American English as a whole.

In the next chapter, I examine the second major strand of evidence in support of the grammaticalization of *you guys* in PdAE: gradual changes in its morphosyntactic constituency and their relationship to frequency of usage.

5. The role of chunking in the development of *you guys*

As outlined in Section 1.2, chunking is a reducing effect of frequency characterized by the erosion of constituency boundaries in repeatedly sequential items in discourse. Of particular interest in this investigation is the morphosyntactic boundary between *you* and *guys*. If we find evidence that the [PRO (AP) AppNP] construction has been collapsing in the case of *you guys*, then we have further confirmation that it is becoming increasingly grammatical, as chunking is a common correlate of frequency gains in grammaticalization (Bybee and Thompson 2000).

I propose three layers of measurement to assess *you guys*'s degree of unithood throughout the 20th century: adjacency, association, and univerbation. The first two measures are adapted roughly from Torres Cacoullos 1999 and the latter from Hopper and Traugott 2003.

The adjacency measure tracks changes in the proportion of items that occur sequentially (*you guys*) versus those with intervening morphemes (*you AP guys*). The second assessment of unithood, association, refers to changes in the proportion of uses with a particular type of item (*guys* as opposed to *men*, *fellows*, and so on) co-occurring with another item (here *you*); an increase in this proportion suggests that a special lexical relationship is being forged between the particular type and its collocate (between *guys* and *you*). The third and final measure of reduced constituency involves univerbation, generally defined as the treatment of an erstwhile phrasal item as a single free morpheme; in more advanced stages, this may involve a free morpheme, or word, becoming bound (Hopper and Traugott 2003: §6, at 'morphologization'). The form *mother-in-laws*, for instance, reflects greater univerbation than *mothers in-law*, the plural marker attaching in

the former case as it would to any applicable word-level stem; in the latter case /-z/ interrupts *mother* and *in-law*, hindering their potential to be processed as a chunk and allowing for an alternative categorization as ‘nominal + embedded modifying phrase’.

The univerbation measure can be tailored to the particular case of *you guys* by exploring (a) the frequency of a newer genitive pattern in which possession is marked on *you guys* rather than on *you* alone (e.g. *you guys’ friends* vs. *your guys friends*), (b) the frequency of *you guys* with contracted auxiliaries compared to that of other 2p variants, in light of the fact that monomorphemic personal PRO subjects comprise the overwhelming majority of subjects with contracted auxiliaries, and (c) the upwelling of orthographically fused *you guys* variants in recent written texts.

In this section, I evaluate changes in *you guys*’s adjacency, association, and univerbation to uncover evidence that it has been increasingly processed as a chunk.

5.1 Changes in the adjacency of *you* and *guys*

Turning to the first measure of unithood, adjacency, one telling indication of *you guys*’s restructuring in American English lies in its changes in usage with intervening APs throughout the 20th century, as exemplified by (22) and (23):

(22) I’ve understood everything *you two guys* were talking about (COHA, *It Pays Advertise*, 1914).

(23) Now *you new guys* get in back of these seasoned troops here (COHA, *Viet Rock*, 1966).

Table 15 provides the token frequencies of all [*you AP guys*] uses in the COHA, organized by decade:

Variant	Token frequency									
	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000
<i>you two guys</i>	1	2	3	4	1	2	3	1		
<i>you damn guys</i>			2							
<i>you four guys</i>				1		1				
<i>you little guys</i>									2	
<i>you new guys</i>						1	1			
<i>you other guys</i>					1		1			
<i>you smart guys</i>			2							
<i>you young guys</i>									2	
<i>you wise guys</i>				1						
<i>you white guys</i>									1	
<i>you TV guys</i>								1		
<i>you truthful guys</i>			1							
<i>you three guys</i>				1						
<i>you stupid guys</i>							1			
<i>you spy guys</i>								1		
<i>you shorter guys</i>							1			
<i>you same guys</i>									1	
<i>you old-time guys</i>							1			
<i>you old guys</i>							1			
<i>you navy guys</i>						1				
<i>you mid-life guys</i>									1	
<i>you medical guys</i>						1				
<i>you married guys</i>									1	
<i>you history guys</i>									1	
<i>you fucking guys</i>								1		
<i>you fitzgerald guys</i>							1			
<i>you film guys</i>							1			
<i>you dumb guys</i>							1			
<i>you D.C. guys</i>										1
<i>you business guys</i>							1			
<i>you black guys</i>								1		
<i>you Bilagaana guys</i>									1	
Total <i>you AP guys</i> tokens	1	2	8	7	2	6	13	5	10	1
Total combined <i>you guys</i> & <i>you AP guys</i> tokens	7	18	70	109	127	188	168	210	355	479
Percent <i>you AP guys</i>	14.3	11.1	11.4	6.4	1.6	3.2	7.7	2.4	2.8	0.2

Table 15: Variants of [*you AP guys*] in the COHA, 20th century

You two guys is the most frequent instance of the [you AP guys] construction in all periods of the COHA except the 1990s and 2000s, when *you little guys* and *you young guys* otherwise assume this position. Although there are more [you AP guys] tokens overall in the latter decades of the COHA (20 for 1910-1959 compared to 35 for 1960-2009), there is a substantial decrease in the proportion of non-adjacent *you guys* tokens among all [you (AP) guys] tokens. This general decline is displayed in Figure 9, based on the bottom row of Table 15:

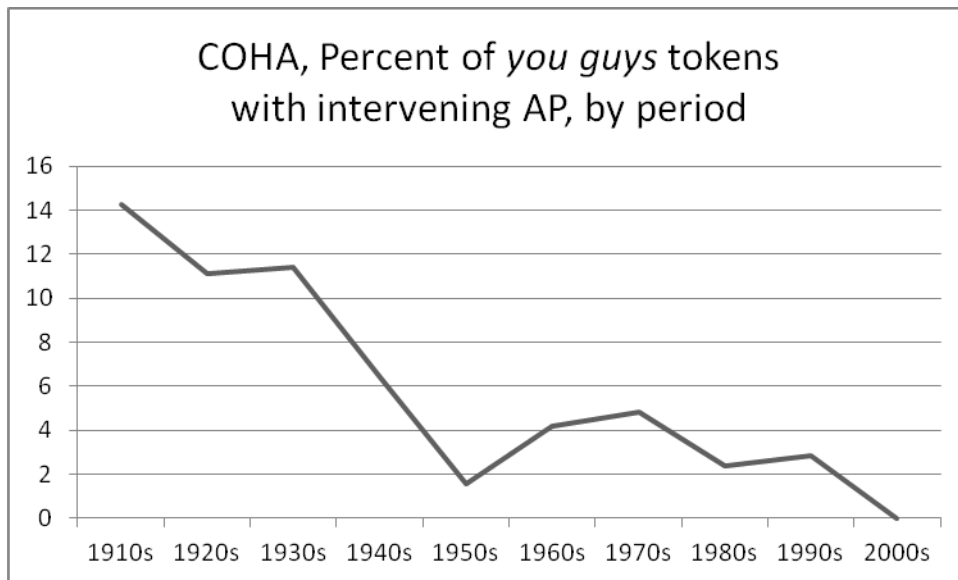


Figure 9: The proportional decline of *you AP guys* usage, COHA

As one would anticipate were chunking underway, the correlation between the percentage of *you AP guys* tokens and decade in Figure 9 turns out to be negative; this relationship is strong, moreover, and achieves statistical significance at $p < .05$ ($p = .003$, $\tau = -.8$, two-tailed). Thus it is extremely unlikely that such a dependency is due to chance.

Figure 9 implies that *you* and *guys* have become proportionally more adjacent throughout the 20th century and thereby more susceptible to chunking. While it is tenable to some extent that *you guys* has retained its lexical associations to the [PRO (AP)

AppNP] construction (Table 15), *you guys*'s alternative categorization as a reduced construction has become more strongly reinforced over time. The reduced construction can be symbolized as [*you*_{PRO} *guys*_{AppNP}], with a pronominal slot restricted to *you*, a gradually weakened AP slot, and a nominal slot specific to *you guys*. The [*you*_{PRO} *guys*_{AppNP}] use therefore represents a historical intermediary between the [PRO (AP) AppNP] construction and the [*you guys*]_{PRO} unit.⁴ In Chapter 6, I advance the notion that the most recent stage in the chunking of *you guys*—its unithood as [*you guys*]_{PRO}—has coincided with the loss of lexical restrictions to masculinity.

The evidence of increasing unithood in Figure 9 adds further refinement to the account of *you guys*'s lexical basis in Chapter 3, for it is not, strictly speaking, the [PRO (AP) AppNP] construction in which *you guys* pronominalized in the latter part of the 20th century, but the newer, intermediary [*you*_{PRO} *guys*_{AppNP}] use whose constituency had been steadily collapsing under the pressures of grammaticalization. In this regard grammaticalization not only occurs in particular constructions (Bybee et al. 1994, Traugott 2003), but in exceptional instances of those constructions in consecutive stages of a form's development.

The co-existence of an older [*you* (AP) *guys*] use and a newer, increasingly chunked pronominal unit represents divergence, a well-known characteristic of grammaticalization (Hopper 1991). Hopper 1991: 22 explains, “When a lexical form

⁴The intermediary construction is perhaps best illustrated by examples in the early to middle part of the 20th century in which *you guys* and *you* are interchanged, implying that *you guys* has become more aligned with the categorization of a PRO unit, but in which *guys* nevertheless is strongly associated with its original masculine appositive function: “From time to time Sorenson would turn to warn Winters and the men around him. ‘Pipe down,’ he would say. ‘Take it easy.’ Winters said, ‘I don’t give a damn. Let him fire me. But *you guys*. *You* going to let him get away with it?” (COHA, Land of Plenty, 1934). A second example: “Take it easy, *you guys*! *You*’ll get your dough! Don’t crowd! Just line up...line up! (Syd herds the group [of men] into a rough line...) (COHA, Bad and Beautiful, 1952). See also examples of “*you guys and you* AP NPs” (e.g. [44] below), where intervening items are avoided for *you guys* yet not for other *you*-appositives.

undergoes grammaticalization...the original form may remain as an autonomous lexical element and undergo the same changes as ordinary lexical items.” The development of *you guys* throughout the 20th century has been marked by a gradual split, in this sense, between its lexical categorization as [PRO (AP) AppNP] and [*you guys*]_{PRO}.

The outcome of the gradual dissociation of the [*you guys*]_{PRO} from the appositive construction can be found quite plainly in the contemporary spoken data, where non-adjacent usage constitutes but a fraction of overall *you guys* tokens. This brings us to Section 5.1.1.

5.1.1 The adjacency of *you* and *guys* in contemporary speech

A similarly minute segment of non-adjacent *you guys* tokens can be found in the spoken section of the COCA, supporting the COHA result that today *you guys* is most often processed as a chunk. Table 16 parallels the previous table by showing the token frequencies of all non-adjacent *you guys* uses in the spoken section of the COCA from 1990 to 2012:

Variant	Token Frequency
<i>you two guys</i>	32
<i>you media guys</i>	2
<i>you press guys</i>	2
<i>you three guys</i>	1
<i>you radio guys</i>	1
<i>you sports guys</i>	1
<i>you poor guys</i>	1
<i>you old guys</i>	1
<i>you oil guys</i>	1
<i>you five guys</i>	1
<i>you defense guys</i>	1
<i>you black guys</i>	1
<i>you automobile guys</i>	1
<i>you AFL-CIO guys</i>	1
Total <i>you X guys</i> tokens	47
Total combined <i>you guys</i> and <i>you X guys</i> tokens	7,267
Percent <i>you X guys</i>	0.7

Table 16: Variants of [*you AP guys*] in the COCA, all periods

Akin to the written data for the 1990s and 2000s, in speech fewer than 1% of all *you guys* tokens involve non-adjacency. Thus in both communicative modes, *you guys* follows the newer pattern lacking an AP far more than the older, more loosely structured one.

The less restrictive morphosyntax of the older pattern is perhaps best illustrated by the incorporation of APs with coordinated adjectives, as in (24) and (25) from the COCA:

(24) But beware, *you social, economic and political climbers*. Don't try to impress Jesus (COCA, Christ Century, 2006).

(25) He waved his arm...as if it proved my guilt. "*You high and mighty know-it-alls!* You think we don't love our family like normal people, you think our kids don't matter, but they do! They matter to us!" (COCA, Bones, 1990).

Many such examples can be found in the COHA as well: *you reckless and rambling boys, you proud and free Americans, you professional and business men, you fine and generous gods, you clever and infallible members, you fair and tender ladies, and you fat and greasy citizens.*⁵ In these examples we encounter the decidedly freer constituency structure of the [PRO (AP) AppNP] construction, in contrast to the large majority of *you guys* uses that have otherwise become rigidified as a sequence.

The traditional account of *you guys* as a synchronically generated instance of the [PRO (AP) AppNP] construction (Postal 1966/1969, Delorme and Dougherty 1972) cannot explain its gradual structural rigidification and increasing unithood within this construction. This shortcoming of the generative perspective underscores the need for an alternative diachronic usage-based account to *you guys*'s constituency.

In the ensuing section, I determine whether the trend toward increasing adjacency is particular to *you guys* rather than being a broader feature of the *you*-appositive class.

5.1.2 Changes in adjacency among other *you*-appositives

A comparison to other frequent instances of the appositive construction establishes the fact that the increasing adjacency of *you* and *guys* is not a more widespread property within the class of *you*-AppNPs, which would otherwise undermine the claim that *you guys* is diverging from this class morphosyntactically.

The development of *you men* and *you fellows* illustrates this point aptly. Since *you men* and *you fellows* were the most frequent instances of the appositive construction prior to the attestation of *you guys* (Table 9) and to a large extent overlapped semantically at

⁵ For continuity I am omitting here examples of [PRO AP...AP AppNP] construction with *we* and *us*.

the beginning of the 20th century, they were in principle as likely to grammaticalize in PdAE.

As Figures 10-11 reveal, however, the proportions of [*you AP men*] and [*you AP fellows*] usage do not change substantially over time, unlike those of [*you AP guys*]:

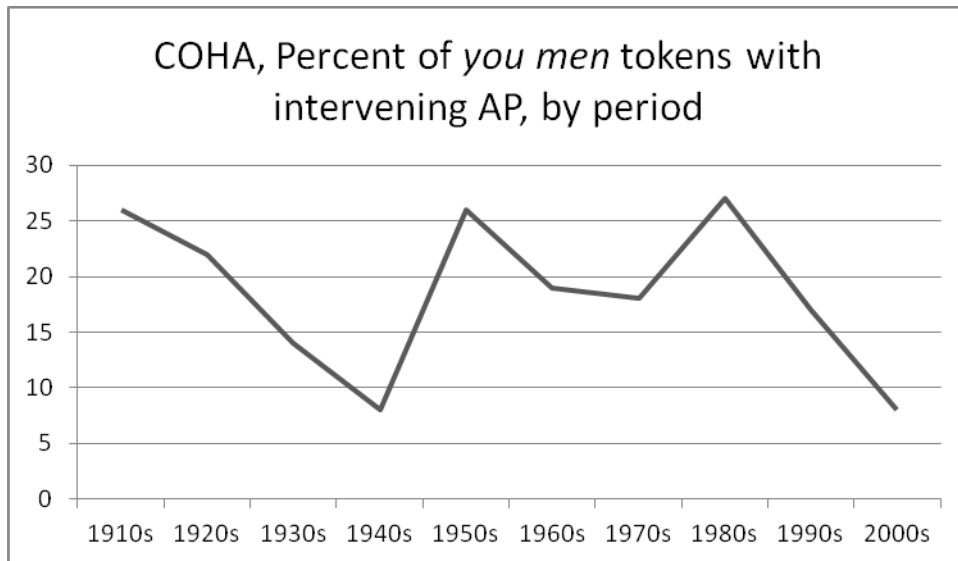


Figure 10: The pattern of [*you AP men*] usage, COHA (Token frequencies: 1910s: 21; 1920s: 17; 1930s: 12; 1940s: 7; 1950s: 25; 1960s: 10; 1970s: 12; 1980s: 15; 1990s: 4; 2000s: 1; All: 124)

Contrary to *you guys*, the negative correlation between the percentage of [*you AP men*] usage and the decade is weak and fails to reach statistical significance at $p < .05$ ($p = .27$, $\tau = -.25$, two-tailed).

You fellows proceeds along a similar path in the COHA, as Figure 11 demonstrates:

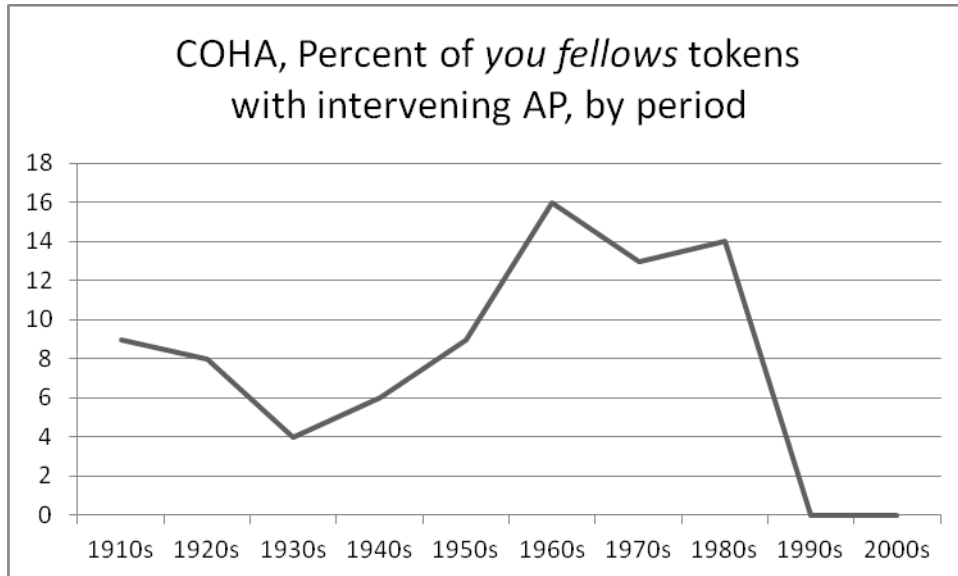


Figure 11: The pattern of [*you AP fellows*] usage, COHA (Token frequencies: 1910s: 14; 1920s: 13; 1930s: 4; 1940s: 2; 1950s: 6; 1960s: 8; 1970s: 3; 1980s: 2; 1990s: 0; 2000s: 0; All: 52)

A Kendall's *tau* test likewise falls short of establishing a dependency between the proportion of [*you AP fellows*] usage and the decade in Figure 11 ($p=.79$, $\tau=-.07$, two-tailed). Unlike *you guys*, semantically similar instances of the appositive construction like *you men* and *you fellows* have failed to make significant gains in adjacency.

Even the more frequent 2p variant *you all* appears to be on a distinct path with respect to changes in adjacency. In the COHA, the following [*you AP all*] uses can be found: *you uncertain all*, *you staid all*, *you sorry all*, *you savvy all*, *you mad all*, *you jumpy all*, *you happy all*, *you half-starving all*, *you crazy all*, *you comfortable all*, and *you anxious all*. Over time, these [*you AP all*] uses have not proportionally declined in any meaningful way, differently from *you guys*:

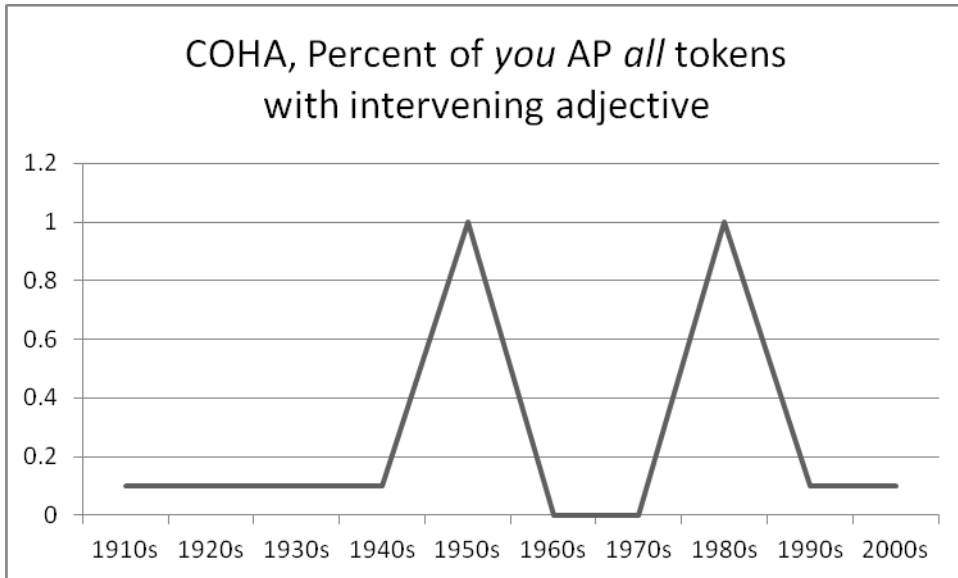


Figure 12: The pattern of [*you AP all*] usage, COHA (Token frequencies: 1910s: 1; 1920s: 2; 1930s: 1; 1940s: 1; 1950s: 5; 1960s: 0; 1970s: 0; 1980s: 4; 1990s: 1; 2000s: 2; All: 17)

A test for dependency establishes this fact, as it produces an extremely weak and insignificant correlation ($p=.92$, $\tau=.03$, two-tailed). To the extent that there has not been a notable change in the incorporation of APs for *you all*, its constituency structure corresponds to the appositive construction more closely than that of *you guys*.

Figure 12 must be qualified, however, in view of the increasing usage of *y'all/ya'll*, which clearly shows concatenation in its orthography. While the portion of [*you AP all*] usage has not decreased in any significant way, the frequency of *y'all/ya'll* has been increasing markedly in the COHA, much akin to *you guys*. Refer to Figure 13:

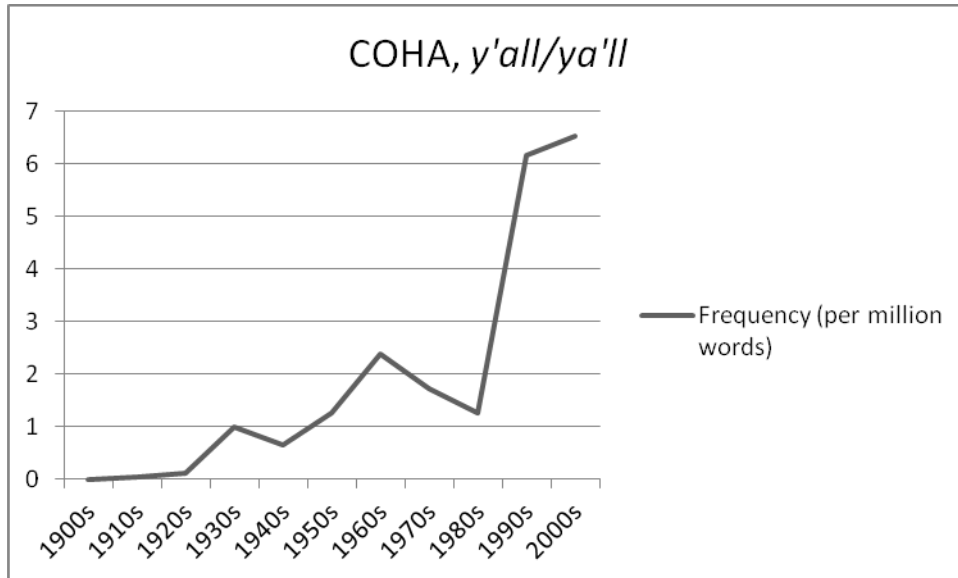


Figure 13: The frequency of *y'all/ya'll*, excluding *you all* and *you-all*, COHA (Token frequencies: 1910s: 1; 1920s: 3; 1930s: 24; 1940s: 23; 1950s: 34; 1960s: 58; 1970s: 42; 1980s: 33; 1990s: 176; 2000s: 228; All: 622)

The dependency between the frequency and decade of *y'all/ya'll* usage is strong and statistically significant (where $p < .05$, $p = .0001$, $\tau = .8$, two-tailed), similar to the finding for *you guys* (Figure 1).

The increase in the frequency of *y'all/ya'll* and simultaneous flatness in the frequency of *you all* (Figure 5) corroborate the finding of Lipski 1993, Maynor 2000, and Tillery et al. 2000 that the former has been diverging from latter in PdAE, and thus from the [PRO (AP) AppNP] construction alongside *you guys*. While these scholars disagree about the origins of *y'all/ya'll*—whether it derives from an African American creole (Lipski 1993), from more ordinary reductive processes (Maynor 2000 and Tillery et al. 2000), or from both—they nevertheless concur that the form differs functionally as well as structurally from *you all*. The reduced variant is more casual, more specifically used by Southerners to reinforce group membership, and more particularly associated with youth registers (ibid.). In terms of grammaticalization theory, *y'all/ya'll* therefore represents another special instance of the appositive construction in which *you* and the

appositive NP (here a quantifying PRO) have formed a processing unit with distinct semantic and pragmatic functions. Though it is outside the scope of the present study, a grammaticalization analysis would in this light be an insightful addition to the existing body of work on *y'all/ya'll*.

5.1.3 Summary and discussion

Figures 9-12 indicate that while *you guys* has been gaining in unithood throughout the 20th century by way of increasing adjacency, other (more ordinary) instances of the [PRO (AP) AppNP] construction have been structurally conservative, retaining the optional AP slot. For this reason such uses as *you men*, *you fellows*, and in some manner *you all* (i.e. excluding *y'all/ya'll*) can be classified as “more lexical” than *you guys* in PdAE. Not only do *you men*, *you fellows*, and *you all* accept intervening lexical material (APs and coordinated APs), but more convincingly, this morphosyntactic property has not changed significantly over time.

Meanwhile, adjacent *you guys* usage has been on the rise in American English. Being more susceptible to chunking as a result, *you guys* has undergone more dramatic changes in constituency than other *you*-appositives, having gradually shifted in its categorization from [PRO (AP) AppNP] to [*you*_{PRO} *guys*_{AppNP}] to [*you guys*]_{PRO}. Resembling the divergence of *y'all/ya'll* from *you all* (again, see Lipski 1993, Maynor 2000, and Tillery et al. 2000), *you guys* has gained in autonomy with respect to its more lexical basis in the appositive construction.

5.2 The increasing association of *you* and *guys* in the appositive construction

Another indication that *you guys* has been increasingly processed as a chunk emerges from assessments of *you*'s association to *guys* in plural address, as opposed to other common NPs that participate in the appositive construction. To this end Table 17 charts usage with *you*-NP (involving a non-*guys* plural common noun) versus that of *you guys* across each decade of the COHA. For brevity, a number of individual *you*-NP (non-*guys*) variants have been collapsed into the category "Other (non-*guys*)" in the table:

Variant	Token Frequency									
	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s
<i>you boys</i>	144	111	121	118	102	81	102	37	110	91
<i>you fellows</i>	167	165	102	67	63	43	19	14	14	4
<i>you men</i>	56	91	67	86	63	57	36	33	19	23
<i>you folks</i>	50	103	67	58	49	43	27	38	32	21
<i>you girls</i>	113	73	51	24	33	21	34	12	31	32
<i>you gentlemen</i>	38	50	48	41	57	27	40	27	16	11
<i>you Americans</i>	20	9	25	34	33	37	32	25	5	16
<i>you kids</i>	13	20	46	17	22	20	41	19	45	41
<i>you children</i>	17	60	26	18	36	15	9	7	23	8
<i>you women</i>	24	41	17	13	16	5	12	5	10	3
<i>you ladies</i>	19	14	19	11	13	15	16	16	10	6
Other (non- <i>guys</i>)	122	118	105	101	112	107	88	82	66	55
Total <i>you</i> NP (non- <i>guys</i>)	783	855	694	588	599	471	456	315	381	311
<i>you guys</i>	6	16	62	102	125	182	255	205	345	478
Total <i>you</i> NP (non- <i>guys</i>) & <i>you guys</i>	789	871	756	690	724	653	711	520	726	789
Percent <i>you guys</i>	1	2	8	15	17	28	36	39	48	61

Table 17: The proportion of *you guys* tokens among all *you*-NP (non-*guys* plural common noun) tokens used as addressives, by decade, COHA

The positive statistical relationship between the percent of *you guys* tokens and the decade in Table 17 (bottom row) proves to be strong and highly significant at $p < .05$

($p = .0001$, $\tau = .96$, two-tailed), providing compelling evidence that *you* has become more

closely associated with *guys* in the appositive construction. This result substantiates the *you guys*'s increasing unithood along the association measure.

The steady growth in the portion of *you guys* tokens relative to other *you*-NPs suggests that the *you*-appositive construction has become more specialized, as it were, for *you guys*. Hopper 1991: 22 characterizes specialization as a reduction in the variety of formal choices in a class occurring when specific forms of the class become more frequent, semantically generalize, and increasingly occupy contexts once filled by other forms of the class. While *you guys* has become more frequent (Chapter 4) and semantically generalized (Chapter 6), it has increasingly occupied contexts in which other *you*-appositives—for instance, *you folks/people* in mixed-gender address or *you girls/gals/women/ladies* in feminine-exclusive address—instead might have appeared (see Sections 6.3-6.5 for examples). In light of my primary purpose of explicating the grammatical progress of *you guys*, it is important to note that specialization, along with layering (Section 4.4.1), divergence (Section 5.1), decategorialization (Section 5.4.1), and persistence (Section 6.1), can be taken as indicators that grammaticalization is underway (Hopper 1991).

5.2.1 Further evidence of association: the COCA spoken data

In addition to evidence in the COHA of *guys*'s increasing association to *you* and the related specialization of the *you*-appositive construction for *you guys*, the COCA spoken data yield a high degree of association between *you* and *guys* in the form of relevance rankings among plural common nouns that follow *you*. This ranking is based on mutual information scores of collocates, or multi-word sequences (Jurafsky et al. 2001), representing the probability that two or more words will co-occur in a corpus

given their individual frequencies and rate of co-occurrence. Table 18 reveals that *guys* is the most relevant collocate after *you* among the class of plural common nouns:

Relevance ranking	Collocate	Token frequency	Percent of all <i>you</i> - NP (plural common noun) collocates	Mutual information score
1	<i>guys</i>	7220	31.74	4.21
2	<i>fellows</i>	29	6.58	1.94
3	<i>updates</i>	18	5.31	1.63
4	<i>nuts</i>	53	4.08	1.25
5	<i>folks</i>	334	2.91	0.77
6	<i>liberals</i>	54	2.42	0.5
7	<i>gentleman</i>	115	1.82	0.09
8	<i>ladies</i>	95	1.77	0.04
9	<i>clues</i>	13	1.2	-0.52
10	<i>notes</i>	35	1.11	-0.62

Table 18: The relevance ranking of NP (plural common noun) collocates after *you*, COCA spoken

Guys comprises nearly 32% of all *you*-NP (plural common noun) collocates in the COCA spoken section, with a mutual information score of 4.21. If we ignore the requirement that *you*'s collocate be a plural common noun, *guys* in fact turns out to be one of the most relevant words occurring after *you* more generally in the COCA.

Relevance ranking	Collocate	Token frequency	Percent of all 'you X' collocates	Mutual information score
1	<i>betcha</i>	56	84.85	5.63
2	<i>know-*</i>	560	70.71	5.37
3	<i>know</i>	244572	60.88	5.15
4	<i>'re</i>	122185	34.19	4.32
5	<i>mention</i>	3157	32.54	4.25
6	<i>foresee</i>	79	32.38	4.24
7	<i>guys</i>	7234	31.8	4.21
8	<i>solemnly</i>	45	29.8	4.12
9	<i>guessed</i>	83	29.23	4.09
10	<i>cannot</i>	21	28.77	4.07

Table 19: The relevance ranking of one-word (or one-gram) collocates after *you*, COCA spoken (*=*know* with an appended hyphen in transcription, usually indicating a pause)

In Table 19, *guys* ranks seventh highest in relevance to *you* in relation to all one-word or one-gram collocates. Thus not only is *guys* closely associated with *you* within the *you*-appositive construction, but also more generally among all *you*-X tokens. *Guys*'s high degree of relevance to *you* in the COCA represents the outcome, or at least the most recent phase, of diachronic changes in the lexical association of *you* and *guys* (Table 17). Due to these changes, *guys*'s lexical representation is now more tightly woven with that of *you*, highlighting the unithood of *you guys* in PdAE.

5.3 The univerbation of *you guys*

A final measure of *you guys*'s unithood relates to its univerbation, or usage as a single free morpheme. In this section I first address the usage of *you guys* with possessive markers, and then turn to its usage with contracted auxiliaries. I finally discuss reduced orthographic variants of *you guys* that have surfaced in recent written texts.

5.3.1 *You guys* versus other *you*-appositives in genitive constructions

Two particular genitive constructions are of interest in the first inquiry: (1) the [*your guys*' NP] construction, in which *your* refers to the same set of participants as *guys* and possessive suffixes occur on *your* and *guys* alike in a sort of agreement relation, and (2) the [*you guys*' NP] construction, in which (similarly) *you* refers to the same set of participants as *guys* but (dissimilarly) lacks the Old English genitive case marker, /-1/.

The first [*your guys*' NP] pattern is shown in (28) and (29), taken from the COHA and COCA, respectively:

- (28) “Heard you was pulling offers to join broadcast booths. Even maybe manage.” “Yeah well,” Sprissel grinned, “it’d need to be a good goddamn offer. To take me away from cleaning out all *your guys’ pockets* week after week” (COHA, Southern Review, 2006).
- (29) Mark Goodin, I never heard you or any Republican strategist when the Teamsters were scratching *your guys’ backs* ever say this was a great calamity. You guys loved the Teamsters... (COCA, CNN Crossfire, 1997).

In (28), retired baseball player Steve Sprissel chides his fellow poker partners in his response to an inquiry into his professional ambitions. In this case *your* and *guys* refer to the same set of participants, the poker partners present in the scene, and possession is marked redundantly on *your* and *guys*. In (29), likewise, the backs that are referred to in the utterance belong to the referents of *your guys* as a unified set of participants rather than to the referents of *guys* alone. In a distinct construction instantiated by *your enemies’ ships*, for example, *your* and *enemies* refer to two different sets of participants, and consequently the ships belong to the referents of *enemies* alone.

The second, alternative pattern for marking possession, the [*you guys’* NP] construction, is illustrated by (30) and (31):

- (30) “I wish I had *you guys’ metabolism*,” he had muttered earlier. “Then I could be skinny too” (COHA, Good as Gold, 1979).
- (31) *You guys’ sister* is getting married tonight (COHA, Ploughshares, 2005).

Most importantly in an analysis of unithood, *you guys* functions as the possessor in (30) and (31) without structural interruption from the genitive /-ɪ/ inflection, thereby enabling the chunking of *you* and *guys*.

In the COHA, the majority of 2p *guys* genitives follow the second [*you guys*' NP] pattern for marking possession. Table 21 offers frequency data for the two *guys* genitive uses from that corpus:

Use	Examples (COHA)	Token Frequency
<i>your guys</i> ' NP	<i>your guys' pockets</i>	1
<i>you guys</i> ' NP	<i>you guys' fault, you guys' outline, you guys' metabolism, you guys' sister</i>	4

Table 21: The token frequency of *you guys* genitive uses, COHA (all periods)

Of the five *guys* genitive uses in this corpus, four, or 80%, adhere to the later [*you* NP -z/-əz/-s NP] pattern. Additionally, 33% of all tokens in the corpus instantiating this pattern are *you guys*' NP (4/12), such that it is the most frequent instance of the newer possessive construction. These results are somewhat speculative, however, given the low token frequencies in Table 21. It becomes necessary in this light to examine the COCA to determine if the distinctive preference for *you guys*' holds there as well.

Supporting such a claim in a corpus with spoken data is more challenging, as the possibility exists that apostrophes have occasionally been overlooked in transcription. For some speakers, moreover, the possessive marker is masked by the articulation of the /z/ in *guys* (e.g. [jə ɡaɪz fɹɛnz] for *you guys' friends* instead of [jə ɡaɪzəz fɹɛnz]). One potential solution would be to exclude the spoken section of the COCA, though this strategy effectually recreates the problem of low overall token frequencies; only 3 examples of each construction can be found in the non-spoken COCA.

To resolve this matter, the spoken section has been admitted into the analysis, but only after the expanded contexts of all *your guys*-NP and *you guys*-NP tokens have been more closely checked to confirm a genitive reading—that is, regardless of whether an

apostrophe appears in transcription. Table 22 provides frequency counts for the two *you guys* genitive uses in the COCA (1990-2009):

Use	Examples (COCA)	Token Frequency
<i>your guys</i> ' NP	<i>your guys' stuff, your guys' pockets, your guys' hands, your guys' decline, your guys' comments, your guys' backs, your guys' administration, your guys names [sic]</i>	8
<i>you guys</i> ' NP	<i>you guys' candidates, you guys' sister, you guys' music, you guys' money, you guys' litmus test, you guys' involvement, you guys' help, you guys' friends, you guys' differences, you guys' last two CDs, you guys' powerful presence, you guys opinion [sic]</i>	12

Table 22: The token frequency of *you guys* genitive uses, COCA (1990-2009)

In the COCA, the preference for *you guys*' is less pronounced, though the majority of *guys* genitives still follow this pattern (12/20, or 60%).

The bias toward the genitive pattern with *you* noted in Tables 21 and 22 provides a preliminary indication of *you guys*'s univerbation. For speakers more typically mark possession on the entire *you guys* sequence as though it were a word-level stem, thereby eschewing the Old English /-ɪ/ suffix. Where used, of course, this suffix has the effect of interrupting the *you guys* string, or inhibiting its chunking. The present section demonstrates, then, the development and predominance of a genitive construction in which *you* and *guys* can remain adjacent and as such, gradually form a unit.

Additionally, *guys* is the most frequent NP-possessor in the *you* genitive pattern, constituting 71% (12/17) of all such uses in the COCA (e.g. *you boys' names, you peoples' positions, or you ladies' workday*). This fact suggests that [*you guys*' NP] has not only become more frequent than [*your guys*' NP], as discussed above, but also has

become the most prominent instance of the [*you* NP (Possessor)] NP (Possessee)] construction in the lexicons of American English speakers.

The preliminary clues of *you guys*'s univerbation that manifest in genitive contexts (Tables 21 and 22) are further substantiated by the form's coalescence with contracted auxiliaries and by its emerging orthographic variants, matters that are taken up in the following sections.

5.3.2 *You guys* versus other *you*-appositives with contracted auxiliaries

An additional source of evidence toward the univerbation of *you guys* can be found in its usage with contracted auxiliaries, namely *'re*, *'d*, and *'ll*, as in (32)-(34):

(32) So, *you guys're* still going through with that? (COHA, Shape Things, 2001).

(33) *You guys'd* be safer coming with us (COHA, Land Dead, 2005).

(34) Geez, the things *you guys'll* do for twenty bucks (COHA, Fantasy Sci Fi, 2001).

Table 23 shows the frequency of *you guys* with contracted auxiliaries in the COHA:

Use	Token Frequency
<i>you guys'll</i>	6
<i>you guys're</i>	6
<i>you guys'd</i>	3
Total	15

Table 23: The token frequency of *you guys* with contracted AUX, COHA (all periods)

Though the overall frequency of *you guys* with contracted auxiliaries is somewhat low in the COHA, the tokens in Table 23 span 13 different texts with distinct authors, suggesting that the usage is not idiosyncratic.

The written portion of the COCA adds further weight to this claim; for the present comparison the spoken section of the COCA was excluded to mitigate against phonetic

ambiguity between *are/re* (and in some cases between *will'll* where the [w] is weakened but not fully elided), leading to potential inconsistencies in transcription:

Use	Token Frequency
<i>you guys'll</i>	8
<i>you guys'd</i>	6
<i>you guys're</i>	2
Total	16

Table 24: *You guys* uses with contracted AUX, COCA written (all periods)

While the frequency ranking of such *you guys* uses differs slightly in the COCA, with *you guys'd* being more frequent in this corpus, *you guys's* usage with contracted auxiliaries is clearly not an artifact of a single author or corpus, instead reflecting a more general pattern in the language.

In light of the manner in which auxiliaries typically contract in English (Krug 1998), uses like those in Tables 23 and 24 reflect the morphosyntactic reduction of *you guys*. In English, auxiliaries contract far more often with monomorphemic rather than periphrastic subjects, especially with monomorphemic pronouns given their high string frequency with auxiliaries (ibid.). The COHA data for *'ll* (contracted *will*) demonstrate this tendency robustly. In a randomized sample of 100 tokens from the 2000s, the following morphemes precede *'ll*, in descending order of frequency:

Morpheme	Token Frequency
<i>I</i>	44
<i>you</i>	20
<i>we</i>	13
<i>they</i>	8
<i>He</i>	6
<i>that</i>	3
<i>she</i>	2
<i>it</i>	2
<i>there</i>	1
<i>Zella</i>	1

Table 25: The frequency of morphemes with contracted 'll (COHA, 2000s)

Of the 100 'll tokens in the sample, 99, or 99%, involve a monomorphemic pronoun.

With a 95% confidence interval for this sample proportion being 1.95 (corrected for a population size of 23,241, or all 'll tokens from the 2000s), we can be reasonably certain that the overall corpus proportion of 'll tokens with monomorphemic pronouns falls in the 97-100% range.

In the written section of the COCA, the same basic pattern is encountered; again, the spoken section of the COCA has been ignored in this analysis to avoid potential inconsistencies in the transcription of 'll.

Morpheme	Token Frequency
<i>I</i>	32
<i>we</i>	24
<i>they</i>	21
<i>you</i>	10
<i>he</i>	5
<i>it</i>	5
<i>she</i>	2
<i>there</i>	1

Table 26: The frequency of morphemes with contracted 'll (COCA, 2000s)

In the randomized sample of 100 *'ll* tokens from the COCA in the same period (2000s), all 100, in fact, involve monomorphemic PRO subjects. These data make it plain that *you guys* is mimicking a pronominal unit rather than a periphrastic constituent when it enters auxiliary constructions like those of (30)-(32).

To emphasize this theme, the usage of *you guys* and other *you*-appositives with contracted auxiliaries can be directly compared. Table 27 therefore lists all contracted *'ll*, *'re*, and *'d* tokens preceded by *you all*, *y'all*, *you men*, and *you fellows* from the COHA, to be contrasted to Table 23; only tokens from the 1910s-2000s have been counted to permit a valid comparison, as *you guys* is first attested in the COHA in the 1910s.

Use	Token frequency
<i>you all'll</i>	2
<i>you all're</i>	0
<i>you all'd</i>	0
<i>y'all'll</i>	0
<i>y'all're</i>	0
<i>y'all'd</i>	1
<i>you men'll</i>	1
<i>you men're</i>	0
<i>you men'd</i>	0
<i>you fellows'll</i>	0
<i>you fellows're</i>	0
<i>you fellows'd</i>	0
Total	4

Table 27: *You all*, *y'all*, *you men*, and *you fellows* uses with contracted auxiliaries (COHA, 1910s-2000s)

Whereas 15 *you guys* uses with contracted auxiliaries can be found in the COHA between 1910 and 2009 (Table 23), only four such tokens can be found for *you all*, *y'all*, *you men*, and *you fellows* combined.⁶ *You guys* is nearly four times more frequent than other *you-*

⁶ Phonological factors are likely at work here as well, especially in the case of *you all/y'all will*, in which the [l]-[l] sequence may be avoided in certain utterances.

appositives when accompanied by a contracted auxiliary. In constructions with contracted auxiliaries, then, we encounter confirmatory evidence of *you guys*'s treatment as a single morpheme (univerbation), in the form of its unique semblance to monomorphemic PRO subjects.

5.3.3 A diachronic look at *you guys*'s univerbation

A final point to be made in this section pertains to the timing of *you guy*'s usage with contracted auxiliaries. The question remains: When exactly did the most dramatic changes in *you guys*'s treatment as a pronominal unit in auxiliary constructions occur? To this end, Figure 15 parcels *you guys*'s usage with contracted auxiliaries into the distinct decades of the COHA:

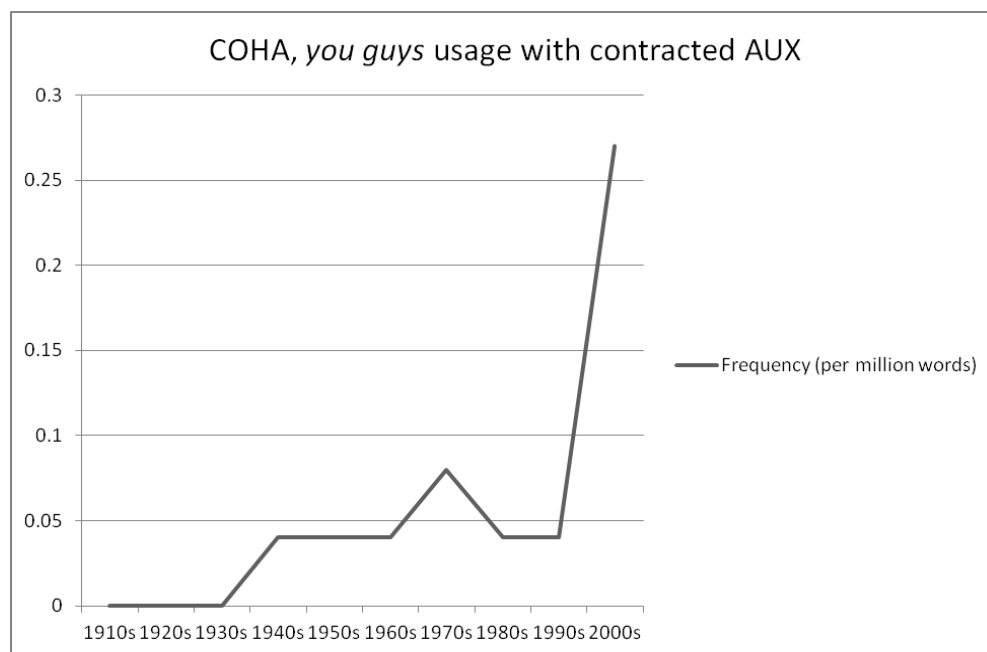


Figure 15: *You guys* usage with contracted auxiliaries (COHA)

You guys first began to be used with contracted auxiliaries in the 1940s, experienced another spike in such usage in the 1970s, and then increased most sharply in frequency in the 1990s. The graph therefore highlights three pivotal periods in *you guys*'s growing

resemblance to monomorphemic pronouns in auxiliary constructions. Not coincidentally, these are the same three periods in which *you guys* undergoes the most dramatic semantic changes as well (Chapter 6).

The positive correlation between frequency and decade of usage in Figure 15 is moderate-to-strong and achieves significance at $p < .05$ ($p = .006$, $\tau = .74$), suggesting that the changes in *you guys*'s usage with contracted auxiliaries have not occurred at random. Instead, as I assert, these changes are due *you guys*'s increasing univerbation in PdAE. A remaining indication of *you guys*'s univerbation involves its recent orthographic variants, explored in Section 5.3.4.

5.3.4 Orthography as a reflection of univerbation

Perhaps the strongest strand of evidence toward the univerbation of *you guys* lies in its orthographic variation. While the orthography of *y'all* more noticeably represents its treatment as a word than that of *you guys*, the latter form has begun to reveal its own history of chunking in its many spellings: *you-guys*, *ya guys*, *ya'guys*, *y'guys*, and *yguys*. The following example from the 1940s portion of the COHA depicts the oldest of the variants, *you-guys*:

- (35) “Look – *you-guys* – ” he panted. “Spread out – along the down-canyonside. – Say – who ain't here?” (COHA, Fire, 1948).

More recently, Maynor 2000 and Tillery et al. 2000 use this hyphenated *you-guys* variant in their manuscripts to note its concatenation.

A much newer orthographic variant, *ya guys/ya'guys*, likewise exhibits the structural reduction of the appositive construction:

(36) That is the first time since 1900 that the Reds have had two one-hit shutouts thrown their way in succession. Good on *ya guys* for rolling back the calendar. (“Tale of Two Games” 2013).

Though no such tokens appear in the COHA or the COCA written section, scores of *ya guys* or related *ya’guys* forms can be found on the internet. A Google search of the *ya guys* string produces nearly two million hits, only 7.2-20.8% of which, in my estimation, are false positives (e.g. “YA guys,” an abbreviation for ‘young adult guys’ and “I’ll meet ya, guys,” where *guys* is used vocatively after *ya*).⁷ This estimation is based on an analysis of 100 randomly selected search results, of which 14 were false positives, yielding a 95% confidence interval of 6.8 (based on the sample size and proportion, corrected for an actual population size of 1,760,000 on July 6, 2013). Conservatively, then, 79.2% of all *ya guys* search results in Google are *ya guys* or *ya’guys* pronominal matches, implying 1,393,920 individual tokens overall.

Among the most visibly reduced orthographic variants of *you guys* is *y’guys*, found in the COCA written section in the 1990s:

(37) I shuddered as it occurred to me that my young son may grow up saying, “Hey *y’guys*, anybody seen my ice skates?” instead of “Hey y’all, anybody seen my lizard-skin ropers?” (COCA, Houston Chronicle, 1997).

In the Google database, a related *yguys* variant (lacking the apostrophe) is also encountered, and is in fact more easily countable due to its continuity and resultant dissimilitude from the false positive “Y Guys,” the name of a popular band.

⁷ Because Google does not record punctuation, a search for *ya guys* produces examples of both *ya guys* and *ya’guys*, among false positives.

Of 31,000 *yguys* search results in Google, 18.3-35.7% are valid *yguys* pronominal matches based on my estimation. As with *ya guys/ya'guys*, in this analysis I counted the number of matches out of 100 randomly selected *yguys* hits, finding 27 pronominal tokens total, or 27% (discounting false positives such as “YGuys,” where the space had been omitted in the popular band’s name). I then computed a 95% confidence interval for this sample proportion, in light of the sample size ($n=100$) and overall population ($N=31,000$ on July 6, 2013). As the confidence interval that emerged from this process was 8.7, the most cautious assessment of the frequency of the *yguys* pronominal in the Google search results is 5,673, or 18.3% of 31,000.

These alternative orthographic variants of *you guys* clearly reflect its increasing univerbation, or representation as a single morpheme, thereby compounding the evidence outlined in Sections 5.3.1 and 5.3.2. As the unithood of *you guys* becomes reinforced in the American English lexicon through changes in adjacency, association, and univerbation, we should expect the higher incidence of orthographically truncated *you guys* variants in American texts. While *you-guys*, *ya guys*, *ya'guys*, *y'guys*, and *yguys* are far rarer in the corpora than *y'all*, the frequency of such variants on the internet is a harbinger, in my mind, of more widespread usage to come.

5.4 Summary and discussion

Given the growing interdependence of *you* and *guys* via changes in adjacency, association, and univerbation, *you guys* has been gradually recategorized as an integrated pronominal form as opposed to a multi-part construction. The treatment of *you guys* as a grammatical unit departs from its traditional classification as an instance of the more lexical [PRO (AP) AppNP] construction (Postal 1966/1969, Delorme and Dougherty

1972), which involves three separate slots. The gradualness of *you guys*'s structural reduction, moreover, indicates that its morphosyntactic properties cannot be adequately explained from a synchronic perspective (ibid.). The corpus data establish the plausibility of an intermediary [*you*_{PRO} *guys*_{AppNP}] construction with a steadily weakened AP slot throughout the 20th century.

More recent characterizations of *you guys* (e.g. Howe 1996, Wales 2004, De Vogelaer 2007) are similarly misleading. By emphasizing *you guys*'s lexical character—for example its status as a “lexical compound” (Howe 1996)—these scholars overlook the gradual structural changes within the [PRO (AP) AppNP] construction that have occurred for *you guys* as it has grammaticalized.

One way to better understand *you guys*'s restructuring is to situate these developments within a broader discussion of ‘analyzability’ and ‘schematicity’ (Langacker 1987). Langacker 1987: 292 defines analyzability as the “recognition of the contribution that each component makes to the composite conceptualization.” In other words, it is the increasing or decreasing ability to lexically parse the various morphemes that comprise a construction, including their constituency, as opposed to processing them holistically. Schematicity refers to the corresponding property of constructions that are lexically analyzable. More specifically, it is the increasing or decreasing incorporation of different types of morphemes in a particular slot in a construction, or the incorporation of more or fewer slots altogether. Thus constructions that incorporate a greater number of different types of morphemes in a greater number of slots are said to be “more schematic” than ones with fewer types/slots.

The history of *you guys* in American English exemplifies these properties of grammatical systems poignantly. Due to increases in adjacency, association, and univerbation (Sections 5.1-5.3), the constituency boundary between *you* and *guys* has become less analyzable to speakers. As a result, *you guys* has become more lexically dissociated from the [*you* (AP) AppNP] construction, which embodies two separate constituency boundaries in the case of AP usage. The [*you* (AP) AppNP] construction has, meanwhile, become less schematic in the special case of *you guys* (and *y'all* for that matter) by gradually losing its AP slot. The [*you* (AP) AppNP] construction has become less schematic in a different way as well. As Section 5.2 demonstrates, *you* has become increasingly associated with *guys* in this construction, or with one particular type of plural common noun.

Langacker's contribution to the field is an important one insofar as analyzability and schematicity offer linguists a "lens" through which to view the integration of the lexicon and grammar of a language. The development of *you guys* illustrates this point well. As *you guys* has become less analyzable and schematic, it has gradually departed from its representation as the more lexical [*you* (AP) *guys*] appositive use, and in doing so has inched closer to its present-day status as a pronominal unit. Approaches to language that entail harder distinctions between the lexicon and grammar (e.g. Chomsky and Halle 1968) cannot readily accommodate such a gradual change from less to more grammatical.

The decreasing analyzability and schematicity of *you guys* suggest that it had become a "prefab," or prefabricated utterance (Erman and Warren 2000, Wray and Perkins 2000), by the middle part of the 20th century. Wray and Perkins 2000: 1 define

prefabs as sequences of words that are “stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar.” One implication of *you guys*’s gains in adjacency, association, and univerbation is it was becoming a “ready-made” processing chunk as it advanced toward pronominalization. As Bybee and Torres Cacoullos 2009 report, a common property of grammaticalization is a higher degree of unithood in prefabs relative to other instances of constructions undergoing change; thus the prefabs tend to grammaticalize sooner and more quickly than the other instances. This is indeed what we find with *you guys*, which revealed greater unithood with respect to other *you*-appositives during its transition from construction to form.

5.4.1 Interdependence of chunking and frequency in *you guys*’s development

To conclude this chapter, it is important to establish a relationship between chunking and frequency in the development of *you guys*, thereby linking Chapters 4 and 5 and underscoring the importance of usage in language change. For this purpose a Kendall’s *tau* test can be used to verify positive correlations between (a) the adjacency of *you guys* (Figure 9) and its frequency (Figure 1), (b) the association of *you* and *guys* (Table 17) and the frequency of *you guys* (Figure 1), and (c) the univerbation of *you guys* (Figure 15) and its frequency (Figure 1). Because *you guys* is grammaticalizing, such a test should uncover intimate connections between its degree of chunking and its rate of occurrence throughout the 20th century. Table 28 reveals the results of this inquiry:

Correlations	τ	p
Adjacency, frequency	.71	.005*
Association, frequency	.96	.0001*
Univerbation, frequency	.74	.006*

Table 28: Correlations between unithood measures and frequency for *you guys* (*=significant at $p < .05$)

Along all three measures, *you guys*'s level of unithood is indeed closely tied to its token frequency, a finding that is quite typical of constructions undergoing grammaticalization and that further supports the usage-based approach of this dissertation.

The culmination of *you guys*'s chunking within the special [*you*_{PRO} *guys*_{AppNP}] construction is its conceptual representation as [*you guys*]_{PRO}. This change exemplifies 'decategorialization', or the loss of morphosyntactic properties in grammaticalization, which Hopper 1991 considers to be a defining trait of the process. In the following chapter, I argue that the [*you guys*]_{PRO} unit emerges most clearly in the corpus data once its masculine meaning undergoes bleaching. This brings us to a discussion of the semantic generalization of *you guys*, which like chunking, is a predictable reducing effect of token frequency in grammaticalization.

6. The semantic generalization of *you guys* and its social dynamics

Many American English speakers can recall a time when *you guys* was used almost exclusively for men, and if used to refer to women, sounded jarring for its gender insensitivity. Given such a perception, *Washington Post* “Etiquette” columnist Judith Martin once quipped to her mixed-gender audience, “The language of respect is fading out of use everywhere....*you guys* have a problem with this?” (cited in Rios 2004, emphasis added). Martin was subtly criticizing the newer mixed-gender and feminine-exclusive meanings of *you guys* for their sexually discriminatory character, reiterating objections to the generic usage of *he*, *man*, and many other masculine-origin terms raised by feminist scholars (e.g. Bate 1978, Baron 1986, Cameron 1998, Eckert and McConnell-Ginet 2003).

In view of the negative attitude about *you guys* encapsulated in Martin’s comment, it is somewhat perplexing that the form has not been explored in any great depth in the feminist literature. Two scholars, Bate 1978: 145 and Eckert and McConnell-Ginet 2003: 69, make brief note of the gender-neutrality of *you guys*, but do so matter-of-factly and only in passing. *You guys*’s lack of thorough scrutiny in this body of research is likely due (in part) to the dissipation of negative social attitudes regarding its non-masculine usage as it has semantically generalized. Unlike the *he* and *man* generics, moreover, the newer non-masculine meanings of *you guys* were pragmatically strengthened in contexts of social intimacy, as I show in Section 6.3-6.4, which aided in mitigating against its negative perceptions. This scenario illustrates the occasional tension in language change between sociolinguistic attitudes, which can be a limiting force

against a form's frequency if sufficiently negative, and the processes of grammaticalization, which correspond to higher frequency of usage.

The case of *you guys* in this sense parallels the development of other English expressions with discriminatory histories on which contemporary speakers rarely pause to reflect. While the term *vandal*, for instance, has retained a general negative connotation, its more specific racist undertones have become obfuscated from today's speakers (< Latin *vandalus* 'a member of an ancient Germanic tribe [derogatory]', OED). The terms *barbarian* and *gyp* similarly derive from discriminatory histories (< Greek *bárbaros* 'foreigner [derogatory]' and Latin *gipcyan* 'Egyptian', OED), yet many speakers today use these forms uncognizantly. And how many speakers describing their child or pet as a "cute little bugger," to give a final example, have been aware of *bugger*'s roots in the Latin term *bulgarus*, indicating a member of a Bulgarian tribe that the Romans stereotyped as heathens and sodomites (OED)? The development of such expressions in English is analogous to *you guys* to the extent that they exhibit the gradual weakening of lexical connections to an older, more socially biased meaning. This chapter proceeds from the vantage point that habituation, and what often follows in the specific case of language change, semantic generalization, underlie *you guys*'s increasing social acceptance in non-masculine address. Semantic generalization has gradually eroded its restriction to masculinity and subsequent ties to gender bias in the newer uses. Chapter 6 also discusses the further influence of enrichment—that is, associative processes cueing on contiguity between *you guys* and its high-frequency contexts of usage, on the conventionalization of *you guys* as a non-masculine addressive.

6.1 The origin and conservation of *you guys*'s masculine semantics

Assuming that Ade's usage in (13) represents the initial entry of *you guys* into American English, the phrase appears roughly 30 years after the term *guy*, meaning 'man, fellow', is itself first attested in the United States (OED, at *guy*). Thus a key precursor of *guys*'s role in the *you*-appositive construction was the prior establishment of *guy* as part of the American English lexicon.

According to the OED (at *guy*, n.₂), *guy* originally had pejorative connotations in British English, meaning 'an effigy burned on Guy Fawkes day' (< Guy Fawkes, a British dissident who attempted the assassination of King James in 1605). *Guy*, however, began to lose such connotations as the purpose of Guy Fawkes day grew more disconnected from the foiled assassination. The OED clarifies that children's creation of effigies on Guy Fawkes day became more focused on the collection of money for fireworks than on the defamation of Fawkes or the celebration of royal might (at *guy*, n.₂); in many respects this development is therefore reminiscent of the de-Christianization of Christmas in many parts of the world.

Due to the growing disconnection between Guy Fawkes day and its original context, the appearance of "guy" effigies became more diverse, as the depiction of Guy Fawkes in particular grew less common. In the following example from 1825, roughly two centuries after Fawkes' assassination attempt, the author comments directly on the changing appearance of the "guy" effigy:

- (38) Formerly an old cocked hat was the reigning fashion for a '*Guy*' ...now, however, both hat and mitre have disappeared (OED, W. Hone Every-day Bk: 1430).

Example (38) implies that new “guy” fashions were emerging in the 1800s, according to which “an old cocked hat” and “mitre” had become unnecessary.

Importantly, the association between the “guy” effigy and masculinity (via Fawkes) has nevertheless persevered, even though the effigy’s appearance has grown increasingly multifarious. A look at the succeeding passage in Hone’s text from 1825 demonstrates this point:

(39) [B]oys carry about their “Guy” with no other sentiment or knowledge respecting him, than body-snatchers have of a newly-raised corpse (OED, W. Hone Every-day Bk: 1433).

Shortly following (38) in Hone’s text, this additional statement by the author refers to the newly fashioned effigy as “him.” Example (39) also emphasizes the fact that by the early 1800s, the effigy had been decontextualized from the events and characters of Guy Fawkes’s assassination attempt, as British children were oblivious, according to the author, to the effigy’s historical appearance and function.

The conservation of masculinity and decontextualization of the effigy meaning provide critical segues to the form’s subsequent masculine human usage. In British English, the first masculine human tokens from the 1860s neatly demonstrate this link by spotlighting the unusual attire of the male human referent:

(40) He was such an old *guy* in his dress (OED, 1861, T. Hughes Tom Brown at Oxf.: II. x. 174).

(41) What are you doing there, dressed up in that way like a *guy*? (OED, 1867, Trollope Last Chron. Barset: II. lx. 175).

The extension of *guy* to humans with the meaning of ‘a male of unusual or striking appearance’ is quite sensible given the similarity between “dressing up” an effigy and “dressing up” (oneself) in a particular costume or outfit.

Not coincidentally, then, the ‘male of unusual or striking appearance’ meaning appears in many of the first uses of *guy* in American English as well, which were roughly contemporaneous with examples (40)-(41) from British English:

(42) [S]topping for a moment to consider, he thought, “No, that will never do! Go to her looking like such a *guy*? Nary time (COHA, What Answer?, 1868).

(43) “[A]nyhow he was a fresh young *guy*, with some sort of uniform hat on. He asked me if I didn’t want him to put my bag up in the rack” (COHA, Stepping Heavenward, 1869).

In each sentence, *guy* is associated with a peculiar, or at least eye-catching, style of dress.

Eventually, however, *guy* lost its restriction to the ‘striking dress’ meaning and generalized to other male referents:

(44) He can eat more and talk more and work less than any *guy* ever I see (COHA, Stepping Heavenward, 1869).

In (44) the particular attire of the referent—and indeed the referent itself in light of the modifier *any*—is unspecified. In contrast to (40)-(43), example (44) therefore conveys a distinctively mundane, or “everyman,” quality in a *guy*.

Within a few decades of *guy*’s semantic extension to men in general, the plural form *guys* began to be used appositively with *you* in plural address, functioning to designate male referents in the discourse. The earliest examples of *you guys*—(1), (3), (5), (7), (9), and (11) above—clarify that it was the extended ‘males, fellows’ meaning of

guys, not the more particular apparel meaning, that initially characterized *guys* in the appositive construction.

The lexical association between *you guys* and its initial masculine human semantics has been conserved throughout the 20th century, shown in (45)-(47):

- (45) 1916: [Mr. Jenks is addressing Baffly and Simmy, both male characters] “Just as we was about to call an ambulance, a gentleman in our building came along and reckonised him as young Mr. Tresslyn. Friend of Mr. Dodge’s. That was enough for us. So I brings him around. Now it’s up to *you guys* to look after him (COHA, From the Housetops).
- (46) 1965: Dodge dealers in the East and Midwest recently ran a radio campaign that openly wooed “*you guys* and gals who are bored with Ford” (*Time Magazine* Corpus, Naming Names).
- (47) 2007: You know, Dave, I take a look at that car of yours, and I am amazed that *you guys* and your wife got out of it safely (COCA, NBC Today).

In (46) and (47), the male-exclusive meaning of *you guys* is strengthened by the addition of *and gals* and *and your wife*, respectively. These more recent examples demonstrate *you guys*’s adherence to a common pattern of semantic change in grammaticalization known as “persistence” (Hopper 1991), whereby older meanings of grammaticalizing forms coexist with newer ones, typically until later stages of the grammaticalization process. While there is ample evidence that *you guys* has been grammaticalizing (Chapters 4 and 5, and below), its incomplete semantic change suggests that it is in an intermediary stage of this process.

6.1.1 The role of “tough guy” talk in *you guys*’s masculine semantics

According to the COHA, register played a key role in the development and evolution of *you guys*’s masculine semantics. The term ‘register’ has taken on many different meanings in linguistics (see Biber 1994, 1995, 2006). Following Biber, I use the term generally to cover any speech variety based on the situational context, including its social composition. In the present study the “tough guy” register has been operationalized as talk between soldiers, police, gangsters/criminals, pugilists, or their combatants.

Such usage is illustrated by (48)-(52), from the decades in which *you guys* was most prolific in “tough guy” talk in the corpus:

- (48) 1914: “Yuh better watch out and not turn your backs on him in the dark, none uh *you guys*. I betche he packs a knife. Them kind always does” (COHA, Flying U Ranch).
- (49) 1923: I never said I didn’t kill him. But that ain’t the same as bein’ a regular murderer. What good did it do me to kill him? I didn’t make nothin’ out of it. Answer yes or no! Yes or no, me elbow! There’s some things you can’t answer yes or no. Give me the once-over, *you guys* (COHA, Adding Machine).
- (50) 1930: Listen, boys. There’s a lot of you hoodlums that I never met before so I thought it was about time we got together, seein’ that I’ve been taking the rap for all the suckers *you guys* have been pushin’ around the country lately (COHA, Doorway Hell).
- (51) 1944: Cut it out, *you guys*! Nine-tenths of a war is waiting, and the other tenth is worse (COHA, Storm Operation).

(52) 1950: POLICEMAN: Well – *you guys* got no business in here – city property.

There are signs all over the place (COHA, Asphalt Jungle).

Table 29 compares the proportion of *you guys* usage among “tough guys” to that of all other masculine contexts across three distinct periods of the 20th century, the 1930s-1940s, the 1970s, and the 2000s. These periods were critical in the emergence of *you guys*’s novel meanings, a point which I elaborate on in subsequent sections. 1930s and 1940s tokens have been combined in the register analysis to increase the number of examples for that period. The counts in the table therefore represent all available masculine *you guys* tokens for each respective period except for the 2000s, the numbers for which are based on a randomized sample of 200 examples (hence the inclusion of a confidence interval in the final column only):

1930s-1940s			1970s			2000s			
Referents of <i>you guys</i>	Tokens	% of Referent Group	Referents of <i>you guys</i>	Tokens	% of Referent Group	Referents of <i>you guys</i>	Tokens	% of Referent Group	CI*
Male (“tough guys”)	68	47	Male (“tough guys”)	61	28	Male (“tough guys”)	9	12	3.4
Male (other)	76	53	Male (other)	153	72	Male (other)	66	88	3.4
Total	144	100	Total	214	100	Total	75	100	N/A

Table 29: Proportion of masculine *you guys* usage in “tough guy” talk, by period, COHA (*=95% confidence interval for ‘% of Referent Group’ based on sample proportions and size [n=200], corrected for population size of 478, or all *you guys* tokens in the 2000s)

In the 1930s-1940s, the proportion of *you guys* usage among “tough guys” in the COHA was 47%, yet this number declines to 28% and 12% over the next two periods. We have evidence, then, that *you guys* has been spreading across masculine contexts throughout the 20th century. *You guys*’s semantic history has not only involved movement across genders, but more subtle movement within the masculine category.

Another aspect of *you guys*'s semantics that gradually changed as it left "tough guy" contexts deals with the degree to which its referents were identified in apposition, including their gender. This brings us to a discussion of *you guys*'s newer indefinite-inclusive and mixed-gender definite meanings.

6.2 *You guys*'s indefinite-inclusive semantics

Based on Hawkins earlier work (1978), Lyons 1999 classifies indefinite structures in language along two parameters: identifiability and inclusiveness. On the one hand, indefinites may involve a diminished level of identifiability, or particularization, of referents in the discourse context (i.e. non-specific reference). As a result indefinites lack a gender or number distinction, or both, and typically lack antecedents in the context.

Example (53) illustrates this sense of indefiniteness for the pronoun *you*:

(53) *You* can never do anything well unless you enjoy it (COHA, Fathers Children, 1987).

On the other hand, indefinites can be used to include an entire group of referents, or conversely, to avoid the exclusion of part of that group's members, as in (54):

(54) "I thought all *you* scientists were never satisfied until you had completed your theories with an experiment?" (COHA, Obelists at Sea, 1933).

In this example, *you* stands for the general class of scientists (including the addressee) rather than a smaller subset of scientists or interlocutors present at the moment of discourse.

In the case of *you guys*, an initial sign that its restriction to masculine-specific semantics was beginning to weaken can be seen in its development of indefiniteness in the 1920s:

- (55) 1922: Another thing the promoters propose to do with the Willard and Detripsey fuss is to drop it into the Polo grounds. Not on our life. Better take another guess, *you guys* who think that the fans are going to stand for that stuff (COHA, Chicago Tribune).

In this example, *you guys* exhibits both semantic properties of indefinites: identifiability and inclusiveness. First, the reporter is addressing a hypothetical group in his/her absence and without prior mention in the text, and therefore does not identify a particular set of addressees; it is also not perfectly clear whether the segment of the public being addressed includes women, men, or both. Second, *you guys* in (53) refers to an entire group of readers (“who think that the fans...”) and does not exclude anyone specifically from this group, revealing inclusiveness. In (55) and below, the presence of post-modifiers supports the indefinite-inclusive reading by underscoring the fact that *you guys* by itself underspecifies the class of referents.

The following examples demonstrate the continuation of *you guys*’s indefinite-inclusive semantics throughout the 20th century:

- (56) 1938: “You’re a capitalist, darling. Why do *you guys* in the ruling class let us get away with it?” (COHA, Prodigal Parents).
- (57) 1977: “Jerry, I don’t understand *you guys* who go fooling around with other women. My Joanie is six different kinds of women and that’s enough for one man” (COHA, Hold Me).
- (58) 2002: “He’s not endorsing Vallas for governor. *You guys* in the media have to understand the difference. When I say ‘I hereby endorse Paul Vallas for governor,’ that’s an endorsement (COHA, Chicago Tribune).

In each use, *you guys* fails to identify the gender of its referents in a definitive sense and further signifies a totality of addressees, many of whom are not present at the moment of discourse. Examples (55)-(58) indicate that indefiniteness was not a fleeting aspect of *you guys*'s semantics (i.e. nonce usage), but was in fact a substantive change in its meaning. I return to this point in Section 6.6, where I give diachronic frequency data for all *you guys* meanings, including the indefinite-inclusive category.

You guys's indefinite-inclusive meaning is, in my mind, a telling predecessor of its mixed-gender definite semantics, for this change shows that restrictions to the older masculine meaning were already eroding prior to the emergence of the first mixed-gender definite uses.

6.3 *You guys*'s mixed-gender definite semantics

Once we recognize that *you guys* was being employed in indefinite address with non-specific gender as early as the 1920s, it becomes less surprising that the form's first mixed-gender definite usage occurs in the COHA in a movie script from 1930, *Widow of Chicago*, by Earl Baldwin:

(59) The two gunmen stare sheepishly at each other. Crestfallen, they slowly put away their guns. Mullins and Polly dance into the scene. Both he and Polly are grinning. MULLINS (lightly) Don't get excited, boys. It's all in fun. FIRST GUNMAN (huffed) Yeah? *You guys* got a swell sense of humor. Mullins and Polly laugh. As they dance off PAN on them several feet, then BOOM UP AND REVEAL...FULL SHOT DANCE FLOOR...revealing Polly and Mullins dancing, surrounded by the other couples on the crowded floor (COHA, *Widow of Chicago*).

In the scene depicted in (59), the dancing couple, Mullins (male) and Polly (female), are addressed by the first gunman as *you guys*, providing clear mixed-gender definite usage. Interestingly, the gunman's "tough guy" talk provides a link between the older masculine use and the newer one in the example, suggesting an underlying pragmatic motivation for such an extension. As I show in the following section, this motivation is multifaceted, however, since the referents are also social intimates (a couple).

Though (59) is the only mixed-gender definite use from the 1930s portion of the COHA, in the 1940s such a meaning begins to establish firmer footing in American English. In her 1945 novel *She Never Reached the Top*, for instance, Elma Lobaugh uses *you guys* periodically in gender-inclusive address:

(60) Thought *you guys* were never going to get here. How're you, Jennie? That blonde you liked couldn't come this week end, Jim. Here, let me take those bags (COHA, *She Never Reached*).

In this scene, Lobaugh's characters Jim and Jennie, who are non-romantic friends, are arriving at the vacation rental of Bernard and Delia, where many other companions will soon meet them. Bernard addresses Jim and Jennie with *you guys* as he approaches their car to help them retrieve their bags. As no other male characters are present in the scene, the form conveys an unambiguously gender-neutral meaning.

In another scene from the Lobaugh's novel, friends Peg, Spike, Bernard, and Jennie are eating breakfast when George calls out from the living room as he nurses a hangover:

(61) "How you can eat! Greasy bacon, greasy eggs..." Peg was sharp. "Shut up, George. Don't spoil it for the rest of us. Because you haven't any will power." He

moaned. “Will power. That’s what I need. Will power. If *you guys* were my friends you’d have stopped me last night. You know what I go through (ibid.)

To the extent that *you guys* is not restricted by gender in examples (59)-(61), such usage echoes the form’s indefinite inclusive semantics. Nevertheless the newer mixed-gender uses identify, or set boundaries on, the particular group of interlocutors being addressed, thereby implying definiteness. This difference represents a further development in its semantics.

To summarize, the foregoing sections show that the associative processes of enrichment and semantic generalization have each participated in *you guys*’s semantic progression. Enrichment has played a role by reinforcing the association between *you guys* and “tough guy” contexts, thereby adding to its masculine meaning early on. Section 6.1.1 also identifies the most important register in *you guys*’s initial gain in frequency in American English. Semantic generalization has affected *you guys*, in turn, by contributing to the loss of restrictions to a masculine definite meaning. As *you guys* became more frequent in “tough guy” and other registers in the early part of the 20th century—and thus speakers grew more habituated to its masculine meaning—this meaning gave way to newer indefinite and mixed-gender definite uses in stepwise fashion.

6.3.1 The role of intimacy in *you guys*’s early mixed-gender semantics

Another way in which enrichment in particular has contributed to *you guys*’s semantic history relates to the form’s initial strengthening of associations to contexts of social intimacy, that is, as the newer mixed-gender and feminine-exclusive meanings began to emerge. Hopper and Traugott (2003: 94) clarify that enrichment applies to “the beginnings of grammaticalization...the motivations that permit the process to begin,

rather than its outcomes. There is no doubt that, over time, meanings tend to become weakened during the process of grammaticalization.” The following sections show how this sort of sequential interaction between pragmatic strengthening and semantic generalization has impacted *you guys*’s semantic development.

For the purposes of the present study, ‘intimate’ registers have been defined as those involving a romantic couple, family members, or friends (i.e. excluding casual acquaintances). The following criteria were used in determining which category a particular *you guys* token falls into:

Category	Criterion for <i>you guys</i> referents
Couple	mention of status as couple in text; mention of marriage or dating; mention of romance, physical intimacy, sex, or co-habitation
Family	mention of family, grandfather, grandmother, father, mother, son, daughter, uncle, aunt, or cousin
Friend	mention of status as friends or friendship; mention of shared history (e.g. being roommates, growing up together, going to school together); discussion of personal or private topics (e.g. issues with a marriage partner or family member, sex, emotional problems); mention of partying, having fun, or hanging out together

Table 30: Operationalization of intimacy semantics

Examples (62)-(64) illustrate each of the categories from Table 30:

- (62) Couple: “I [Whitey, a male character] taught her to play chess,” I said. “She whips my ass every time.” “So it’s for real, then. *You guys* are a thing?” Ray Ray asked. Whitey laughed at him. “A thing? A thing?” (COHA, Southern Rev, 2000).
- (63) Family: “Hey, Uncle John. Abuelita [grandmother]. Hey, *you guys*. Where are you going?” (COHA, Boxcar, 2002).
- (64) Friend: “Hey, Ivan, you know we have a big party down here tonight. *You guys* [Ivan and Catherine] should come hang out with us. Big Deke

is throwing it.” “Maybe another night, man. We’re just old friends catching up” (COHA, Pecking Order, 2008).

Table 31 shows the results of this sorting process for the 1930s-1940s, 1970s, and 2000s periods, respectively:

1930s-1940s			1970s			2000s			
Referents of <i>you guys</i>	Tokens	% of Referent Group	Referents of <i>you guys</i>	Tokens	% of Referent Group	Referents of <i>you guys</i>	Tokens	% of Referent Group	CI*
Mixed-gender (family)	0	0	Mixed-gender (family)	4	14	Mixed-gender (family)	12	17	3.9
Mixed-gender (friends)	5	83	Mixed-gender (friends)	15	51	Mixed-gender (friends)	10	14	3.7
Mixed-gender (couple)	1	17	Mixed-gender (couple)	8	28	Mixed-gender (couple)	21	30	4.8
Mixed-gender (other)	0	0	Mixed-gender (other)	2	7	Mixed-gender (other)	27	39	5.2
Total	6	100	Total	29	100	Total	70	100	N/A

Table 31: Proportion of mixed-gender *you guys* usage in intimate registers, by period, COHA (*=95% confidence interval for ‘% of Referent Group’ based on sample proportions and size [n=200], corrected for population size of 478, or all *you guys* tokens in the 2000s)

It becomes immediately apparent in the table that today’s mixed-gender usage of *you guys* reflects a gradual expansion from contexts of friendship (1930s-1940s) to those of family and couples (1970s) and eventually to those involving casual acquaintances and strangers. As the proportion of friendship contexts declines from 83% in the first period to 14% in the final period, the proportions of family, couple, and other (non-intimate) contexts increase. Thus while pragmatic strengthening played an early part in reinforcing *you guys*’s mixed-gender definite meaning within contexts of intimacy, semantic generalization played a later part in eroding the intimacy constraint on this meaning.

Consequently in the 2000s, non-intimates constitute 39% of all mixed-gender definite *you guys* instances in the COHA sample.

As a result, it has become increasingly common for Americans to recruit *you guys* in the address of individuals with whom they are only casually acquainted or with whom they have no acquaintance. In (65), for example, Bill and Chelsea Clinton are addressed as *you guys* while taking questions from a live audience of primary school children during a 1993 ABC television special. One child in attendance, Demetrius, asks the following question about the Clinton's dog, Socks:

(65) Does Socks – who trains him? Is he trained?...Like do *you guys* play with him?

CHELSEA: Yes. DEMETRIUS [To the president]: And also, do you ever have to talk to her [Chelsea] about playing with her [Socks] when she's supposed to do her homework? PRESIDENT CLINTON: Never. She's very good about that
(COCA Spoken, ABC Special).

In a later example from the television news show *20-20* airing in 2010, host Chris Cuomo rebukes interviewees Joseph and Rebecca, who are embroiled in a fiery conflict about which religion to introduce to their child:

(66) It is shameless of you to get a camera crew in and parade your kid as if she were a doll. *You guys* are having at each other and the person who is suffering is your 3-year-old (COCA Spoken, ABC 20-20).

Since the addressees in (65) and (66) have no prior personal relationship with the speaker, these more recent examples embody the historical weakening of *you guys*'s intimacy restriction in mixed-gender address. By juxtaposing examples (62)-(64) and (65)-(66), we witness the historical interplay between pragmatic strengthening and

semantic generalization in the development of *you guys*, or the initial fortification of the form's associations to contexts of intimacy and subsequent weakening of those associations.

A more detailed examination of *you guys*'s contexts of intimacy in Table 31 illustrates the gradual nature of its semantic change. While *you guys* was highest in frequency in mixed-gender friend contexts in the 1930s-1940s (83%), its usage among friends and couples increases in the 1970s and slightly more in 2000s, aiding in the decline of friendship-specific usage (ultimately to 14% in the 2000s). The development of *you guys* has not only been shaped by its loss of restrictions to intimacy, then, but by subtle, piecemeal changes *within* the intimacy semantic that promoted the form's expansion to other intimacy-related contexts (e.g. to families in the 1970s period).

It should be noted, finally, that although the segment of non-intimate (other) mixed-gender usage grows in Table 31, the portion of intimacy usage (friends, couples, and families combined) still constitutes the majority of mixed-gender tokens within each period. In the 2000s, for instance, 63% of all mixed-gender tokens in the sample, or 43 of 70, involve the social cohesion of interlocutors. The intimacy part of *you guys*'s meaning has clearly been conserved in mixed-gender definite address, even though the form is no longer as highly restricted to such contexts due to its further semantic generalization.

In the next section, I discuss *you guys*'s subsequent semantic generalization to female-exclusive contexts, showing how this development is tied to contexts of intimacy. The interaction between semantic generalization and the social composition of the discourse again proves to be critical in an explanation of *you guys*'s novel feminine-exclusive semantics.

6.4 *You guys*'s feminine-exclusive semantics

The intimacy association discussed in Section 6.3.1 provides a “missing link” of sorts between older mixed-gender and newer female-only meanings of *you guys*. Roughly 40 years after the emergence of mixed-gender usage in the COHA, the first female-only uses appear, exemplified by (67) from 1978:

(67) MARTIN. [We're] getting tickets to the ballet for a week from Friday. Do you want to go, Patti?...PATTI A week from Friday. I don't know. Let me ask Phil. ERICA Phil can come with us. MARTIN Only if he pays. PATTI I don't think Phil likes ballet. MARTIN Neither do I and I'm going. ERICA (to Martin) It's good for your soul. PATTI I think I'll pass, Mom. Martin gets up. Kisses Erica and Patti. MARTIN I'll see *you guys* tonight. Have a nice day (COHA, *Unmarried Woman*).

In this scene from the script of Paul Mazursky's film *Unmarried Woman*, Martin addresses Erica and her daughter Patti with *you guys* when no other characters are present in the scene. Such a use is striking not only for its early female-only meaning, but for its similarity to prior mixed-gender definite usage based on the intimacy semantic.

Associations with intimacy, in other words, have aided in mediating the older mixed-gender and (some of) the newer female-only uses. This fact becomes clear by examining changes in the proportion of *you guys* feminine-exclusive uses involving intimacy compared to other feminine-exclusive uses:

1930s-1940s			1970s			2000s			
Referents of <i>you guys</i>	Tokens	% of Referent Group	Referents of <i>you guys</i>	Tokens	% of Referent Group	Referents of <i>you guys</i>	Tokens	% of Referent Group	CI*
Female-only (family)	0	0	Female-only (family)	0	0	Female-only (family)	5	22	4.4
Female-only (friends)	0	0	Female-only (friends)	1	50	Female-only (friends)	13	57	5.2
Female-only (couple)	0	0	Female-only (couple)	0	0	Female-only (couple)	0	0	N/A
Female-only (other)	0	0	Female-only (other)	1	50	Female-only (other)	5	22	4.4
Total	0	0	Total	2	100	Total	23	100	N/A

Table 32: Proportion of feminine-exclusive *you guys* usage in intimate registers, by period, COHA (*=95% confidence interval for '% of Referent Group' based on sample proportions and size [n=200], corrected for population size of 478, or all *you guys* tokens in the 2000s)

Although in the 1970s just one of two female-only *you guys* tokens functions in the address of intimates in the COHA, in the 2000s this proportion climbs to 79% (combining friend and family categories).⁸ In this respect today's feminine-exclusive meaning is at a similar developmental stage as the earlier mixed-gender use from the 1970s, when *you guys* still clustered at high frequencies in intimacy talk (Table 31). Thus the fuller force of semantic generalization has yet to be leveled against the intimacy constraint in feminine-exclusive *you guys* address, as it was eventually in mixed-gender address.

According to Table 32, the intimacy meaning of *you guys* has also spread across types of intimates in feminine-exclusive address, namely from friends to family members. This development neatly parallels the semantic extension of *you guys* in

⁸ *You guys* address of female same-sex couples is unattested in the 2000s COHA sample, hence the exclusion of the couples category in the 79%.

mixed-gender definite address (Table 31). In other words, the emergence of each of *you guys*'s non-male definite meanings reflects the accumulation of small-scale changes in the form's associations to intimacy.

Langacker's notion of 'compositionality' (1987) clarifies the cognitive impact of *you guys*'s weakening restriction to masculinity. Compositionality centers on how predictable the meaning of a multi-word sequence is based on the combined meanings of its parts (ibid.). In the case of *you guys*, the consequence of its semantic generalization has been the decreasing predictability of its meaning based on the combined meanings of *you* '2p addressee' and *guys* 'men, fellows'. The fact that such a change is unique to *you guys* among other *you*-appositive constructions with similar semantic origins implies its recategorization as a more unified (i.e. holistically processed) instance of this construction. The contemporary meanings of *you men* and *you fellows* echo this contrast. Out of 42 *you men* tokens from the 1990s and 2000s periods of the COHA (combined), all 42 encode a masculine meaning; as for *you fellows*, all 18 examples from the same portion of the corpus similarly specify male referents. Thus while *you guys*'s lexical restriction to masculinity has eroded and the form has become less compositional, *you men* and *you fellows* have fully retained this restriction and their higher degree of compositionality.⁹

⁹For brevity I am omitting a discussion of the *guys*-vocative, which also occurs today in mixed-gender definite and feminine-exclusive address. I consider the *guys*-vocative and *you guys*-vocative as separate uses that have developed alongside one another in PdAE, but in very different ways. The former vocative use is nominal and therefore lexical in nature, while the latter, I argue, has become a pronominal unit (i.e. grammatical) via its repetition in the *you*-appositive construction.

6.5 Overview of *you guys*'s non-masculine definite semantics

The influence of the social composition of the discourse on *you guys*'s semantic progression becomes even more palpable by combining mixed-gender and female-only intimate tokens (Tables 31 and 32), in order to track the form's usage among intimates overall. Refer to Table 33:

1930s-1940s			1970s			2000s			
Referents of <i>you guys</i>	Tokens	% of Referent Group	Referents of <i>you guys</i>	Tokens	% of Referent Group	Referents of <i>you guys</i>	Tokens	% of Referent Group	CI*
Non-male (family)	0	0	Non-male (family)	4	13	Non-male (family)	17	18	4.1
Non-male (friends)	5	83	Non-male (friends)	16	52	Non-male (friends)	23	25	4.6
Non-male (couple)	1	17	Non-male (couple)	8	26	Non-male (couple)	21	23	4.5
Sub-total	6	100	Sub-total	28	90	Sub-total	61	66	N/A
Non-male (other)	0	0	Non-male (other)	3	10	Non-male (other)	32	34	5.0
Total	6	100	Total	31	100	Total	93	100	N/A

Table 33: Proportion of non-male (definite) *you guys* usage in intimate registers, by period, COHA (*=95% confidence interval for '% of Referent Group' based on sample proportions and size [n=200], corrected for population size of 478, or all *you guys* tokens in the 2000s)

The sub-total row in the table, which represents all non-male *you guys* tokens with intimate referents for each respective period, confirms two important results in the register analysis. First, contexts of intimacy have been vital in the development of *you guys*'s non-masculine meanings, encompassing 100%, 90%, and 66% of all such meanings across the three periods. This finding underscores the importance of pragmatic motivations in the emergence of *you guys*'s novel gender meanings, as large portions of

the newer uses were consolidated in intimate social contexts.¹⁰ Second, the proportion of intimacy meanings for *you guys* has been declining, falling 44% from the 1930s-1940s period to today, while non-intimate usage has been on the rise. The gradual weakening of the intimacy restriction demonstrates the effect of semantic generalization on earlier, pragmatically enriched components of *you guys*'s meaning.

This latter trend suggests that intimacy is becoming less relevant in the selection of *you guys* for non-male addressees, as example (68) illustrates:

- (68) SONYA: Is it Adriana from Virginia? 5th CALLER [ADRIANA]: Virginia, yes. SONYA Go ahead. 5th CALLER: Hi! How are *you guys* doing? (COCA, CNN Sonya, 1993).

Here *you guys* is part of a ritualistic greeting used in the address of CNN call-in show host Dr. Sonya Friedman and her guest Beverly Johnson, even though the caller has no prior relationship with her interlocutors. This example demonstrates, then, the very sort of non-intimate usage that is on the rise in Table 33. *You guys*'s expansion to non-intimate contexts characterizes its continued semantic progress in contemporary American English, as speakers become more habituated to its lexical restriction to intimacy and in turn more desensitized to this addressive context when selecting *you guys*.

¹⁰The gender of the speaker does not play any clear role in the emergence of *you guys*'s non-masculine meanings. In the 1930s-1940s period, when the mixed-gender definite meaning develops, 50% of the mixed-gender *you guys* tokens have male speakers (2/4). Likewise, in the 1970s, when the feminine-exclusive meaning develops, 50% of the *you guys* feminine-exclusive tokens have male speakers (1/2). These facts point to the role of other social factors in *you guys*'s semantic changes.

6.5.1 Summary and discussion

Lawson 1982: 158 is perhaps the first to identify a tone of “friendliness, camaraderie” in *you guys*, though her evidence for this semantic association is strictly anecdotal. Being written in the early 1980s, moreover, her paper lacks sufficient temporal perspective to inspect the more recent changes in *you guys*’s intimacy meaning. The corpus data in Sections 6.3.1-6.5 more thoroughly substantiate Lawson’s anecdotal finding that intimacy is a critical component of *you guys*’s semantics; these data make it plain, however, that the effect of intimacy on the form’s usage has been attenuated since the 1970s (Table 33).

In light of *you guys*’s decreasing analyzability and compositionality (Chapter 5), there is one additional point in Lawson’s analysis that is arguably problematic. She asserts that *guys* “serves almost as a suffix to *you*....and could be called a ‘register particle’” (1982: 158). That is, *you* serves the normal deictic functions of a pronominal, including, presumably, the redundant expression of plurality, while *guys* encodes the speaker’s stance toward the discourse and its social composition. Other researchers have similarly assumed that *you guys* retains a divisible, two-part structure in PdAE (e.g. Jochnowitz 1982: 69). Given *you guys*’s formal and functional reduction within the *you*-appositive construction, a more reasonable assessment, however, is that it constitutes a processing unit. For this reason *you guys* can be thought to express deictic, grammatical person/number, and pragmatic meanings in a holistic rather than partitive manner.

Intriguingly, the role of intimacy in the development of *you guys* noted by Lawson and expounded upon in the present study mirrors *thou*’s semantics from the late ME and EModE periods (Section 1.1), though *thou*’s intimacy use arose differently via

the expansion of *you* and successive restriction of *thou* to a subset of contexts (Table 3). While their origins differ considerably, *thou* and *you guys* nevertheless overlap in their intimacy-marking function, to this extent exemplifying “renewal” in language change (Hopper and Traugott 2003: 122-4). Of course there are innumerable alternative means for expressing intimacy in language—for example, terms of endearment, statements of affection, tonal changes, and gestures—so it is unnecessary to take the further step and conclude that *you guys* “filled a gap in” the pronominal paradigm of English due to the loss of a T/V distinction. This move would repeat the mistake that English experts have made in oversimplifying the motivation for periphrastic 2p addressives (again, see Mencken 1936, Trudgill and Chambers 1991, Maynor 2000, Hickey 2003, Quinn 2009). The complexity of *you guys*’s pragmatic history in the COHA—from contexts involving “tough guys,” to those of friends, family members, couples, and eventually casual acquaintances and strangers—shows that the form cannot be boiled down to one functional motivation in particular.

In subsequent research it would be instructive to learn whether *you-p*, *you guys*, *you all*, and *y’all* are asymmetrically distributed over contexts of social superior vs. inferior, public vs. private, low emotion vs. heightened emotion, and so on. This line of inquiry would reveal the extent to which the contemporary American English pronominal system recapitulates the T/V system of late ME and EModE. Within the frame of the present study, however, we can at least determine that contexts of social intimacy were critical in the development of *you guys*’s novel meanings, akin to the newer *thou* meanings of late ME and EModE.

The novel, non-male definite meanings of *you guys* have led some scholars to label the form “gender neutral” (Waksler 1995, McLennan 2004). Until the present study, however, the diachronic path of this development was little understood. Chapter 6 provides an explanation as to how exactly such a change has arisen for *you guys*. While the earliest waves of non-male *you guys* usage reveal considerable overlap with particular contexts of intimacy (e.g. among friends), subsequent usage reflects movement across intimacy contexts (e.g. among couples and family members). Also, as *you guys*’s intimacy restriction loosened internally by extending to newer types of intimates, its meaning began to generalize to non-intimates and such newer uses proportionally increased. *You guys*’s generic gender therefore reflects the accumulation of numerous small-scale changes both within and across the form’s newer uses, aptly illustrating the concepts of gradualness and emergence in language change.

6.6 Interdependence of semantic generalization and frequency in *you guys*’s development

In Chapter 5 it was shown that the chunking of *you guys* has been closely tied to its frequency gains throughout the 20th century. The question remains, then, whether *you guys*’s semantic generalization similarly correlates with its increasing frequency of usage. To this end, the present section first defines the proportions of older and newer *you guys* uses within each decade in the COHA (1910s-2000s), and then examines the more general relationship between the proportion of newer uses and the form’s frequency over time (Figure 1).

Table 35 tracks the frequency distribution of *you guys* across its various meanings in the COHA, organized by decade. The table reflects all available *you guys* tokens from

the 1910s through the 1980s, and a randomized sample of 200 tokens per decade for the 1990s and 2000s. For formatting reasons, confidence intervals are given separately and abbreviations are utilized, a key for which follows:

Abbreviation	Key
M	Masculine
MG	mixed gender (definite)
F	feminine-exclusive
I	Indefinite
U	uncodable/ambiguous
TF	token frequency
%	percent of all you guys tokens in period (or in sample)
tot	Total

Table 34: Key to abbreviations in Table 35

	10s		20s		30s		40s		50s		60s		70s		80s		90s		00s	
	TF	%	TF	%	TF	%	TF	%	TF	%	TF	%	TF	%	TF	%	TF	%	TF	%
M	6	100	14	88	53	86	90	88	112	90	176	97	214	84	145	71	113	57	75	38
MG	0	0	0	0	1	2	5	5	2	2	3	2	29	11	34	17	55	28	70	35
F	0	0	0	0	0	0	0	0	0	0	0	0	2	1	7	3	15	8	23	12
I	0	0	1	6	8	13	7	7	9	7	2	1	4	2	12	6	15	8	26	13
U	0	0	1	6	0	0	0	0	2	2	1	1	6	2	7	3	2	1	6	3
tot	6	100	16	100	62	100	102	100	125	100	182	100	255	100	205	100	200	100	200	100

Table 35: Proportions of *you guys* semantic uses, by period, COHA

The proportion of masculine *you guys* uses in the table remains relatively high until the 1970s, which suggests that the initial surge of indefinite and mixed-gender definite usage did not cut into the form's masculine semantics too greatly. Since the 1970s, however, we see the impact of these newer meanings more noticeably as the proportion of male uses gradually declines to 38% (2000s).

As *you guys*'s masculine usage has declined, its non-male definite meanings have conversely been on the rise. Mixed-gender definite usage, for instance, remains relatively low frequency until the 1970s, but then steadily increases to 35% of all *you guys* tokens

in the 2000s; in the 2000s, in fact, mixed-gender definite instances become nearly as frequent as masculine ones, with a marginal difference of 3%.

Feminine-exclusive *you guys* usage has likewise increased since the 1970s, underscoring the importance of this period as a key turning-point in the form's semantic development. From that decade forward, the proportion of *you guys* tokens in female-only address steadily increases to 12% (2000s). This increase is notable even at the lower end of the confidence interval for the sample proportion of feminine-exclusive tokens from the 2000s. Refer to Table 36:

Use	1990s		2000s	
	%	CI	%	CI
M	57	4.5	38	5.1
MG	28	4	35	5.1
F	8	2.4	12	3.4
I	8	2.4	13	3.6
U	1	0.9	3	1.8

Table 36: 95% confidence intervals for 1990s and 2000s '%' from Table 35, given sample proportions and sizes (n=200), corrected for population sizes (N=345 and 478, or all *you guys* tokens from the 1990s and 2000s, respectively)

There is a 95% probability that the overall COHA proportion of feminine-exclusive *you guys* uses from the 2000s falls in the 8.6-15.4% range. Given the lower limit of this interval (8.6%), we can safely conclude that feminine-exclusive tokens are now at least 8.6 times more frequent than they were in the 1970s, when they constituted just 1% of the *you guys* total (Table 35).

Table 36 further demonstrates that the overall COHA proportion of mixed-gender *you guys* usage from the 2000s most probably lies in the 29.9-40.1% range.

Conservatively, then, *you guys* occurs in mixed-gender address 15 times more often today than it did in the 1930s, when it comprised only 2% of all *you guys* tokens. When mixed-gender and feminine-exclusive *you guys* uses are considered together, finally, they make

up no less than 38.5% of all *you guys* tokens in the 2000s section of the COHA (8.6% + 29.9%, reflecting the lower limit of each sample proportion of non-male uses). Non-male definite *you guys* meanings are therefore at least 19 times more frequent today than in the 1930s.

To allow for a more detailed comparison between the proportions of male and non-male definite *you guys* meanings over time, Figure 16 combines the feminine-exclusive and mixed-gender numbers from Table 35:

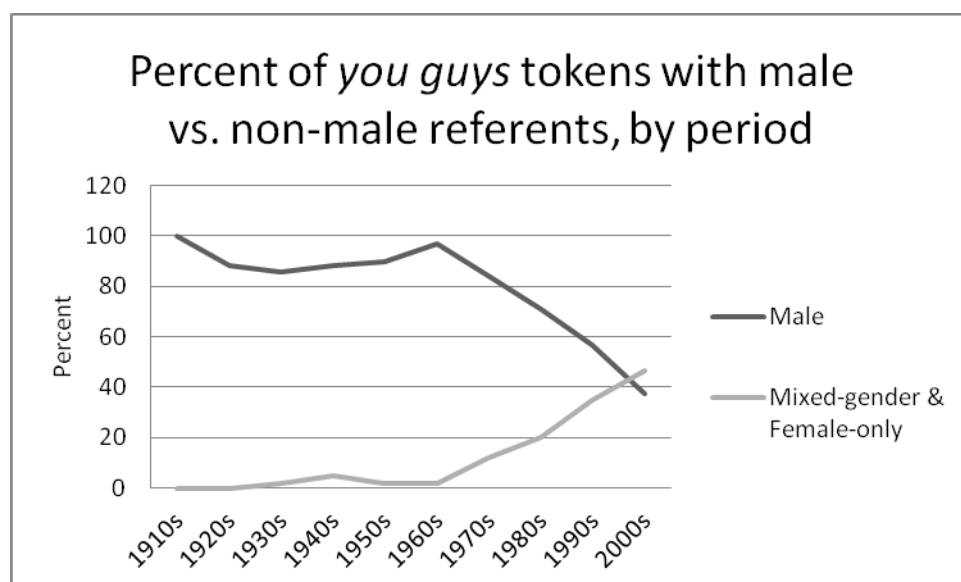


Figure 16: Changes in the proportions of male and non-male (definite) *you guys* usage, by period, COHA

A Kendall's *tau* test reveals a moderate and significant negative correlation between the percentage of *you guys*'s masculine usage and decade in Figure 16 (where $p < .05$, $p = .02$, $\tau = -.02$, two-tailed). Meanwhile the percentage of non-male tokens and decade relate in the opposite direction (i.e. positively); this latter dependency is even stronger and extremely significant at $p < .05$ ($p = .0001$, $\tau = .9$, two-tailed). As a result it is highly unlikely that the relationship between male/non-male *you guys* usage and time in the COHA is due

to chance.¹¹ Instead, as I have shown, this correspondence in the data reflects the accretion of many small-scale changes to *you guys*'s semantics throughout the 20th century, themselves stemming from the multiple interactions of pragmatic strengthening and semantic generalization.

A final matter to be investigated in this section is the more general relationship between the proportion of newer *you guys* uses in Figure 16 and the form's frequency (Figure 1). Given the positive correlation between semantic generalization and frequency that is widely attested in grammaticalization (Section 1.2), an additional piece of evidence that *you guys* has grammaticalized can be found by linking the form's increasing non-male usage to frequency gains.

Indeed, this is the very correspondence we encounter in the COHA, as an analysis of dependency reveals (where $p < .05$, $p = .0001$, $\tau = .9$). Thus the dependency between the proportion of novel *you guys* meanings in Figure 16 and the form's normalized frequency in Figure 1 turns out to be positive and strong at $\tau = .9$, with only a .01% probability of this result being due to chance. In addition to the relationship between chunking and frequency examined in Chapter 5, this result reveals an additional trait of a grammaticalizing form: significant frequency gains associated with greater degrees of semantic generalization, and what follows, decreasing resemblance to earlier, more lexical manifestations.

¹¹ A similar test for correlation between the proportion of indefinite *you guys* tokens and decade leads to insignificant results (where $p < .05$, $p = .2$, $\tau = .3$, two-tailed), suggesting that the proportion of indefinite usage has not shifted considerably over time. The more formative changes to *you guys*'s semantics, then, involve its loss of restrictions to a masculine meaning on the one hand, and its simultaneous reinforcement of non-masculine definite meanings on the other (Figure 16).

6.7 Summary and discussion

As anticipated in an ongoing, diachronic process of language change such as grammaticalization, the semantic progress of *you guys* has thus far been gradual. Over the course of more than a hundred years in American English, there has been a stepwise extension of the form's meaning as successive layers of semantic associations have eroded:

- (*guy*'s association with the 'dandy', or peculiar/striking appearance, meaning)
- *you guys*'s association with the masculine definite meaning in indefinite-inclusive contexts
- *you guys*'s association with "tough guys" in masculine definite contexts
- *you guys*'s association with the masculine definite meaning in intimate non-masculine contexts
- *you guys*'s association with social intimates in non-masculine contexts

The next stage of *you guys*'s semantic development appears to involve the further loss of restrictions to a human-animate meaning.

For instance, in an article from the March 11, 2001 edition of the *Rocky Mountain News* focused on the private life of NBA basketball player George McCloud, McCloud is quoted while addressing his pit bulls with *you guys*:

- (69) "Do *you guys* want a treat?" McCloud asks. He asks in that voice dog owners save for these conversations. They do, and McCloud obliges (BeDan 2001)

In an American-based blog titled *Gracie's New Start*, to give another example, the author contributes the following thread in a discussion about dogs on April 28, 2009:

(70) Just now my mom let the dogs in and she said, “Do *you guys* want a biscuit?” And Murphy started making noises and shaking his head!...He’s a really good dog! (*Gracie’s New Start* 2009).

Non-human animate *you guys* usage can be found in the COHA as well, especially in science-fiction texts involving the address of aliens:

(71) Atop her head, the mousse-encapsulated, balsam-besotted alien seemed laboriously to take cognizance of its altered surroundings...“Shut up! I’m going to make sure *you guys* never mess with my world again!” [Erin says.] Erin felt the Caterpillar strive to regain control of her mind. But she was too strong for it now (COHA, Fantasy Sci Fi, 2001).

In this context *you guys* refers to extraterrestrial Caterpillars with whom the protagonist Erin is engaged in a mental and physical battle. While examples such as (69)-(71) are extremely difficult to find in the COHA and elsewhere on the internet, implying their low token frequency, they nevertheless represent a preliminary stage in *you guys*’s generalization to non-human animates.

You guys’s repeated loss of restrictions to older meanings strengthens the proposal of Chapter 5 that the form has become increasingly autonomous from the [PRO (AP) AppNP] construction. Not only does the AP slot in this construction go unfilled in the overwhelming majority of *you guys* uses, but also *guys* has been increasingly dissociated from its original role of renaming, or designating, male participants in the discourse. The newer indefinite, mixed-gender, feminine-exclusive, and non-human animate uses of *you guys* relinquish this older appositive function. Thus the morphosyntactic and semantic

evidence converge on the conclusion that the *you*-appositive construction has given way to a less compositional pronominal form in the special case of *you guys*.

6.7.1 *You guys*'s semantic change from a cross-linguistic perspective

While *you guys* has undergone a highly particular set of formal and functional changes leading to its pronominalization, its origin in a more loosely structured construction with a masculine human meaning is not entirely unique from a cross-linguistic perspective. Constructions involving nominal terms for humans or human types (e.g. 'person', 'human', 'man', 'guy', 'fellow', 'lord', 'slave') are in fact recurring sources for personal pronominals in the world's languages (Heine and Kuteva 2002, Brown and Levinson 1987). Additionally, constructions involving plural nominal terms for humans (e.g. 'people', 'children') are common sources of plural markers (Heine and Kuteva 2002: 36, 230-1). Within the realm of 2p pronominals in particular, 'human' lexical origins are exemplified by Tok Pisin, with its form *yupela* (< *you fellows*), Spanish, with its unfamiliar/formal form *ustedes* (< *vuestra merced*, 'your majesty'), and Japanese, with its familiar form *kimitachi* (< *kimi* 'lord' + plural suffix). The American English form *you guys* likewise instantiates a grammaticalization path for 2p pronominals that begins with human-type nominals. In this light, *you guys* has always been semantically susceptible to grammaticalization by virtue of its etymological relationship with *guys*.

Though the changes discussed in Chapters 4-6 indicate that *you guys* has advanced considerably along the human-type grammaticalization path, its grammatical progress is as yet partial due to its occasional usage with intervening APs and to the persistence of its masculine meaning. The unfinished weakening of associations to *you*

guys's older lexical properties therefore epitomizes the gradual and gradient nature of autonomy (Bybee and Scheibman 1999, Bybee 2010).

7. Conclusions

Chapters 4-6 detail *you guys*'s morphosyntactic and semantic changes in conjunction with its frequency gains, building a case for its grammaticalization in PdAE. We have confirmation in the COHA data that *you guys* can also serve the range of morphosyntactic functions of *you-p*, and according to the Harvard Survey of North American Dialects, is broadly distributed in areal acceptance along with *you-p*. Given this particular set of results, one can reasonably infer the grammatical status of *you guys* in American English.

Butters 1992: 333 appropriately highlights the paradoxical treatment of *you guys* as a lexical, or sub-grammatical, element (e.g. Smith 1964, Postal 1966/1969, Delorme and Dougherty 1972, Howe 1996, Wales 2004, De Vogelaer 2007): "The newest form, *you guys*, though widespread in the United States...is not often thought of as a pronoun, though increasingly it functions that way in actual usage." In light of the foregoing chapters, Butters is clearly on point. As *you guys* has lost its older, more lexical properties while taking on increasingly grammatical ones, there is no reason to continue to regard it as a lexical construction. Because the processes of language change that have shaped *you guys* have operated gradually, moreover, its characterization as a synchronically generated outcome (Smith 1964, Postal 1966/1969, Delorme and Dougherty 1972) is severely limited, if not altogether misleading.

The alternative diachronic usage-based account of *you guys* developed in the present paper undermines Gramley and Pätzold's assertion (1992: 288) that "to all intents and purposes English has only one second person pronoun, *you*." In American English, what we in fact encounter in 2p address is the co-evolution of *you-p*, *you guys*, *you*

all/y'all, and other more regionally specific pronominal forms, each with varying degrees of grammaticality and inimitable structural, semantic, and pragmatic histories.

This study additionally shows that the motivation for periphrastic 2p structures cannot be reduced to their function in number disambiguation, as pragmatic concerns often overshadow this function. At key stages in the development of *you guys*, for example, its meaning became fine-tuned to the immediate discourse, whether the social context involved “tough guys,” friends, family members, or romantic couples. In addition to number marking, there are countless other discourse needs that speakers in fact fulfill in 2p address, as the usage of *you guys* demonstrates (see also Brown and Gilman 1960, Brown and Levinson 1987, Wales 1983, Mühlhäusler and Harré 1990, and Walker 2003).

Recent psycholinguistic work has established, moreover, that speakers draw from a wide range of paralinguistic resources (e.g. gestures or eye gaze) to spell out *you's* number meaning (Gupta et al. 2007, Frampton et al. 2009). If such paralinguistic means are adequate in number disambiguation, why, then, should periphrastic structures like *you guys* be indispensable for this task?

As the processes of language change that have operated on *you guys* have depended on general cognition and on the form's repetition in particular discourse situations, the current investigation reveals the importance of our experiences with language in shaping its structure. This work also exemplifies the manner in which grammatical outcomes emerge from small-scale changes that “collude” across behavioral domains (Elman 1999). *You guys's* development as a pronominal unit in American English nicely illustrates this collusion by involving numerous minor changes in its processing style (chunking), rate of repetition (frequency), and social context of usage.

References

- Ade, George. 2003/1941. *Stories of Chicago* [Reprint of *Stories of the Street and of the Town*, 1941]. Urbana: University of Illinois.
- Axley, Lowrey. 1926/1927. *You-all, we-all* again. *American Speech* 2: 343.
- Barbieri, Federica. 2008. Patterns of age-based linguistic variation in American English. *Journal of Sociolinguistics* 12.1: 58-88.
- Baron, Dennis. 1986. *Grammar and Gender*. New Haven: Yale University.
- Bate, Barbara. 1978. Nonsexist language use in transition. *Journal of Communication* 28.1: 139-149.
- Bates, Elizabeth, and Judith Goodman. 1999. On the emergence of grammar from the lexicon. In B. MacWhinney (ed.), *Emergence of Language*, 29-79. Mahwah: Lawrence Erlbaum.
- Bauer, Laurie 1983. *English Word-formation*. Cambridge: Cambridge University.
- Bauer, Laurie. 2007. Some grammatical features of New Zealand English. *New Zealand English Journal* 21: 1-25.
- Baugh, Albert. 1978. *History of the English Language*. Englewood Cliffs: Prentice-Hall.
- BeDan, Michael. 2001. Class and struggle: George McCloud exhibits the former; it's the latter that has defined his life. *Rocky Mountain News*: March 11, 2001.
- Biber, Douglas. 1994. An analytical framework for register studies. In D. Biber and E. Finegan (eds.), *Sociolinguistic Perspectives on Register*, 31-55. Oxford: Oxford University.
- Biber, Douglas. 1995. *Dimensions of Register Variation: A Cross-linguistic Comparison*. Cambridge: Cambridge University.

- Biber, Douglas. 2006. *University language: A corpus-based study of spoken and written registers*. Amsterdam: John Benjamins.
- Brinton, Laurel, and Elizabeth C. Traugott. 2005. *Lexicalization and Language Change*. Cambridge: Cambridge University.
- Brown, Penelope, and Stephen Levinson. 1987. *Politeness*. Cambridge: Cambridge University.
- Brown, Roger, and Albert Gilman. 1960. The pronouns of power and solidarity. In T. Sebeok (ed.), *Style in Language*, 253-276. Cambridge: MIT.
- Butters, Ronald. 1992. Grammatical structure. In J. Algeo (ed.), *The Cambridge History of the English Language VI: English in North America*, 325-339. Cambridge: Cambridge University.
- Bybee, Joan L. 1985. *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam/Philadelphia: John Benjamins.
- Bybee, Joan L. 1988. Morphology as lexical organization. In M. Hammond and M. Noonan (eds.), *Theoretical Morphology*, 119-141. San Diego: Academic.
- Bybee, Joan L. 2001. *Phonology and Language Use*. Cambridge: Cambridge University.
- Bybee, Joan L. 2002. Word frequency and context use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change* 14: 21-290. Reprinted in Bybee 2007: 235-264.
- Bybee, Joan L. 2006. From usage to grammar: the mind's response to repetition. *Language* 82: 529-551.
- Bybee, Joan L. 2010. *Language, Usage, and Cognition*. Cambridge: Cambridge University.

- Bybee, Joan L., and Mary A. Brewer. 1980. Explanation in morphophonemics: changes in Provençal and Spanish preterite forms. *Lingua*, 52: 201-242.
- Bybee, Joan L., and William Pagliuca. 1985. Cross-linguistic comparison and the development of grammatical meaning. In J. Fisiak (ed.), *Historical Semantics, Historical Word Formation*, 59-83. The Hague: Mouton.
- Bybee, Joan L., Revere Perkins, and William Pagliuca. 1994. *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World*. Chicago: University of Chicago.
- Bybee, Joan L., and Joanne Scheibman. 1999. The effect of usage on degrees of constituency: the reduction of *don't* in English. *Linguistics*, 37-4: 575-596.
- Bybee, Joan L., and Sandra A. Thompson. 2000. Three frequency effects in syntax. *Berkeley Linguistic Society* 23: 65-85.
- Bybee, Joan L., and Rena Torres Cacoullos. 2009. The role of prefabs in grammaticization: How the particular and the general interact in language change. In Roberta L. Corrigan, Edith A. Moravcsik, Hamid Ouali, and Kathleen Wheatley (eds.), *Formulaic Language 1, Distribution and Historical Change*, 187-217. Amsterdam: John Benjamins.
- Cameron, Deborah. 1998. *The Feminist Critique of Language*. New York: Routledge.
- Campbell, A. 1959. *Old English Grammar*. Oxford: Oxford University.
- Chambers, J. K. 2003. *Sociolinguistic Theory: Linguistic Variation and its Social Significance*. Oxford: Blackwell.
- Chambers, J. K., Trudgill, P. and Schilling-Estes, N. 2002. *Handbook of Language Variation and Change*. Malden: Blackwell.

- Chase, W.G., and H.A. Simon. 1973. Perception in chess. *Cognitive Psychology* 4: 55-81.
- Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.
- Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge: MIT.
- Chomsky, Noam, and Morris Halle. 1968. *The Sound Pattern of English*. Cambridge: MIT.
- Croft, William. 2000. *Explaining Language Change*. Harlow, Essex: Longman.
- Croft, William. 2001. *Radical Construction Grammar*. Oxford: Oxford University.
- Croft, William. 1990/2003. *Typology and Universals*. Cambridge: Cambridge University.
- Croft, William, and Alan Cruse. 2004. *Cognitive Linguistics*. Cambridge: Cambridge University.
- Dahl, Östen. 1985. *Tense and Aspect Systems*. Oxford: Basil Blackwell.
- Delorme, Evelyne, and Ray Dougherty. 1972. Appositive NP constructions. *Foundations of Language* 8: 2-29.
- De Vogelaer, G. 2007. Innovative 2pl.-pronouns in English and Dutch. Flemish Fund for Scientific Research (University of Gent). Unpublished ms.
- Du Bois, John W. 1985. Competing motivations. In J. Haiman (ed.), *Iconicity in Syntax*, 343-365. Amsterdam/Philadelphia: John Benjamins.
- Eckert, Penelope, and Sally McConnell-Ginet. 1999. New generalizations and explanations in language and gender research. *Language in Society* 28.2: 185-201.
- Eckert, Penelope, and Sally McConnell-Ginet. 2003. *Language and Gender*. Cambridge: Cambridge University.
- Elman, Jeff. 1999. The emergence of language: a conspiracy theory. In B. MacWhinney (ed.), *Emergence of Language*, 1-28. Mahwah: Lawrence Erlbaum.

- Erman, Britt, and Beatrice Warren. 2000. The Idiom Principle and the Open Choice Principle. *Text* 20: 29-62.
- Ferber, Edna. 1911. *Dawn O'Hara, The Girl Who Laughed*. New York: Fredrick A. Stokes.
- Fillmore, Charles J. 1971/1997. Lectures on Deixis. Stanford: Stanford University.
- Fillmore, Charles J., Paul Kay, and Mary C. O'Connor. 1988. Regularity and idiomaticity in grammatical constructions. *Language* 64: 501-38.
- Fillmore, Charles J., and Paul Kay. 1994. Grammatical constructions and linguistic generalizations: the What's X Doing to Y? construction. Unpublished ms.
- Finegan, Edward, and John Rickford. 2004. *Language in the USA*. Cambridge: Cambridge University.
- Frampton, M., and R. Fernandez, P. Ehlen, M. Christoudias, T. Darrell, and S. Peters. 2009. Who is "you"? Combining linguistic and gaze features to resolve second-person references in dialogue. In Proceedings of the 12th Conference of the EACL.
- Givón, Talmy. 1973. The time-axis phenomenon. *Language* 49: 890-925.
- Givón, Talmy. 1975. Negation in language: pragmatics, function, ontology. *Working Papers on Language Universals* 18: 59-116.
- Gobet, F., A.J. de Voogt, and J. Retschitzki. 2004. *Moves in Mind: The Psychology of Board Games*. Hove: Psychology.
- Goldberg, Adele E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago.

- Goldberg, Adele E. 2003. Constructions: a new theoretical approach to language. *Trends in Cognitive Science* 7: 219-24.
- Gracie's New Start. 2009. Online blog: post April 27, 2009.
<graciesnewstart.blogspot.com/2009/04/i-have-nothing-to-say.html>.
- Gramley, Stephan, and Kurt-Michael Pätzold. 1992. *A Survey of Modern English*. London: Routledge.
- Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In *Universals of Language*, 73-113. Cambridge: MIT.
- Greenberg, Joseph H. 1966. *Language Universals: With Special Reference to Feature Hierarchies*. The Hague: Mouton.
- Greenberg, Joseph H. 1978. Diachrony, synchrony and language universals. In J. Greenberg, C. Ferguson, and E. Moravcsik (eds.), *Universals of Human Language: Method and Theory*, vol. 1, 61-92. Stanford: Stanford University.
- Gries, Stefan Th. To appear. Useful statistics for corpus linguistics. University of California-Santa Barbara. Unpublished ms.
- Gupta, S., J. Niekvasz, M. Purver, and D. Jurafsky. 2007. Resolving *you* in multi-party dialog. In *Proceedings of the 8th SIGdial Workshop on Discourse and Dialogue*.
- Haiman, John. 1994. Ritualization and the development of language. In W. Pagliuca (ed.), *Perspectives on Grammaticalization*, 3-28. Amsterdam/Philadelphia: John Benjamins.
- Hawkins, John A. 1978. *Definiteness and Indefiniteness*. Atlantic Highland: Humanities.

- Heine, Bernd, and Mechthild Reh. 1984. *Grammaticalization and Reanalysis in African Languages*. Hamburg: Helmut Buske Verlag.
- Heine, Bernd, and Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University.
- Heyd, Theresa. 2010. How you guys doin'? Staged orality and emerging plural address in the television series *Friends*. *American Speech* 85:1.
- Hickey, Raymond. 2004. *Legacies of Colonial English*. Cambridge: Cambridge University Press. 586-620.
- Hills, E. C. 1926/1927. The plural forms of *you*. *American Speech* 2: 133.
- Hopper, Paul J. 1987. Emergent grammar. *Berkeley Linguistic Society* 13: 139-157.
- Hopper, Paul J. 1991. On some principles of grammaticization. In E.C. Traugott and B. Heine (eds.), *Approaches to Grammaticalization* 1, 17-35. Amsterdam: John Benjamins.
- Hopper, Paul J., and Sandra A. Thompson. 1984. The discourse basis for lexical categories. *Language* 60: 4.
- Hopper, Paul J., and Elizabeth Traugott. 2003. *Grammaticalization*. Cambridge: Cambridge University.
- Howe, Stephen. 1996. *The Personal Pronouns in the Germanic Languages*. Berlin: Walter de Gruyter.
- Jochnowitz, G. 1983. Another view of *you guys*. *American Speech*. 58.
- Johnson, Keith. 1997. Speech perception without speaker normalization. In K. Johnson and J.W. Mullennix (eds.), *Talker Variability in Speech Processing*, 145-65. San Diego: Academic.

- Kennedy, A.G. 1915. *The Pronoun of Address in English Literature of the Thirteenth Century*. Stanford: Stanford University.
- Kerswill, P. 2003. Dialect levelling and geographical diffusion in British English. In D. Britain and J. Cheshire (eds.), *Social Dialectology*, 223–243. Amsterdam: Benjamins.
- Krug, Manfred. 1998. String frequency: A cognitive motivating factor in coalescence, language processing, and linguistic change. *Journal of English Linguistics* 26: 286-320.
- Labov, William. 1966. *The social stratification of English in New York City*. Washington D.C.: Center for Applied Linguistics.
- Lakoff, George, and Mark Johnson. 1980. *The Metaphors We Live by*. Chicago: University of Chicago.
- Lakoff, George. 1987. *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago: University of Chicago.
- Langacker, Ronald. 1987. *Foundations of Cognitive Grammar: Theoretical Prerequisites*, vol. 1. Stanford: Stanford University.
- Langacker, Ronald. 1991. *Foundations of Cognitive Grammar: Descriptive Application*, vol. 2. Stanford: Stanford University.
- Larsen-Freeman, Diane. 1997. Chaos/complexity science and second language acquisition. *Applied Linguistics* 18: 141-165.
- Lawson, Sarah. 1982. *Guy and guys*. *American Speech* 57.2: 157-8.

- Lehmann, Christian. 1982. *Thoughts on Grammaticalization: A Programmatic Sketch*, vol. 1. (Arbeiten des Kölner Universalien-Projekts 48). Köln: Universität zu Köln. Institut für Sprachwissenschaft.
- Lindblom, Björn, Peter MacNeilage, and Michael Studdert-Kennedy. 1984. Self-organizing processes and the explanation of language universals. In B. Butterworth, B. Comrie and Ö. Dahl (eds.), *Explanations for Language Universals*, 181-203. Berlin/New York: Walter de Gruyter.
- Lipski, J. 1993. *Y'all* in American English: from black to white, from phrase to pronoun, *English World-wide* 14: 23-56.
- Lobaugh, Elma. 1945. *She Never Reached the Top*. New York: Crime Club.
- Lyons, Christopher. 1999. *Definiteness*. Cambridge: Cambridge University.
- MacWhinney, Brian. 1999. The emergence of language from embodiment. In Brian MacWhinney (ed.), *Emergence of Language*, 213-256. Mahwah: Lawrence Erlbaum.
- Maynor, N. 1996. The pronoun *y'all*: questions and some tentative answers. *Journal of English Linguistics* 24, 288-294.
- Maynor, N. 2000. Battle of the pronouns: *y'all* versus *you-guys*. *American Speech* 75.4: 416-418.
- McLennan, Sean. *Guy, guys, and gender neutrality*. Paper presented at the 40th Chicago Linguistics Society, pp. 211-219, April 15, 2004.
- Mencken, H.L. 1936. *The American Language*. New York: Alfred A. Knopf.
- Miller, George A. 1956. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review* 63: 81-97.

- Mowrey, Richard, and William Pagliuca. 1995. The reductive character of articulatory evolution. *Rivista di Linguistica* 7.1: 37-124.
- Mühlhäusler, Peter, and Rom Harré. 1990. *Pronouns and People: The Linguistic Construction of Social and Personal Identity*. Oxford: Blackwell.
- Murray, Thomas, and Beth Simon. 2008. Colloquial American English: grammatical features. *The Americas and the Carribean*, 401-427.
- Newell, Allen. 1990. *Unified Theories of Cognition*. Cambridge: MIT.
- Nosofsky, Robert M. 1988. Similarity, frequency, and category representations. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 14: 54-65.
- Penelope, Julia. 1990. *Speaking Freely*. New York: Pergamon.
- Pierrehumbert, Janet B. 2001. Exemplar dynamics: word frequency, lenition and contrast. In J. Bybee and P. Hopper (eds.), *Frequency and the Emergence of Linguistic Structure*, 137-58. Amsterdam: John Benjamins.
- Pierrehumbert, Janet B. 2002. Word-specific phonetics. In C. Gussenhoven and N. Warner (eds.), *Laboratory Phonology 7*, 101-39. Berlin: Mouton de Gruyter.
- Postal, Paul. 1966/1969. On so-called 'pronouns' in English. In D. Reibel and S.A. Schane (eds.), *Modern Studies in English*, 201-224. Englewood Cliffs: Prentice-Hall.
- Quinn, Heidi. 2000. Variation in New Zealand English syntax and morphology. *New Zealand English Journal*: 173-197.
- Quinn, Heidi. 2009. Pronoun forms. In Peters et al. (eds.), *Comparative Studies in Australian and New Zealand English*, pp. 31-47. Amsterdam: John Benjamins.

- Rafal, Jeremy. 2009. On the structure of associative plurals: a view from Hawaiian Creole English. Paper presented at the 45th Annual Meeting of the Chicago Linguistics Society, April 2009.
- Richardson, G. 1984. Can *y'all* function as a singular pronoun in Southern dialect? *American Speech* 59: 51-9.
- Rios, Delia. 'You-guys': It riles Miss Manners and other purists, but for most it adds color to language landscape. *Seattle Times*: June 1, 2004.
- Rosenbloom, Joshua, and William Sundstrom. 2003. The decline and rise of interstate migration in the United States: evidence from the IPUMS, 1850-1990. NBER Working Paper 9857.
- Sankoff, Gillian. 1980. *The Social Life of Language*. Philadelphia: University of Pennsylvania.
- Smith, Carlota S. 1964. Determiners and relative clauses in a generative grammar of English. *Language* 40.1: 37-52.
- Stidston, R.O. 1917. *The Use of Ye in the Function of Thou in Fourteenth-Century England*. Stanford: Stanford University.
- Tagliamonte, Sali. 2012. *Variationist Sociolinguistics*. West Sussex: Blackwell.
- Tagliamonte, Sali, and A. D'Arcy. 2009. Adolescence, incrementation, and language change. *Language* 85.1: 58-108.
- "Tale of Two Games." 2013. *My Cincinnati Reds*. Online blog: post April 27, 2013. <blogredmachine.com/2013/04/27/a-tale-of-two-teams/>.
- Talmy, Leonard. 1988. The relation of grammar to cognition, In Brygida Rudzka-Ostyn (ed.), *Topics in Cognitive Linguistics*, 165-205. Amsterdam: John Benjamins.

- Talmy, Leonard. 2000. *Toward a Cognitive Semantics: Concept Structuring Systems*, vol. 1. Cambridge: MIT.
- Tao, H. 2001. Discovering the usual with corpora: the case of *remember*. In R.C. Simpson and J.M. Swales (eds.), *Corpus Linguistics in North America*, 116-144. Ann Arbor: University of Michigan.
- Tillery et al. 2000. The nationalization of a Southernism. *Journal of English Linguistics* 28: 280-294.
- Torres Cacoullos, Rena. 1999. Variation and grammaticalization in progressives: Spanish *-ndo* constructions. *Studies in Language* 23: 25-59.
- Traugott, Elizabeth C. 1972. *A History of English Syntax*. New York: Holt, Rinehart, and Winston.
- Traugott, Elizabeth C. 1982. From propositional to textual to expressive meanings: some semantic-pragmatic aspects of grammaticalization. In W. Lehmann and Y. Malkiel (eds.), *Perspectives on Historical Linguistics*, 245-71. Amsterdam: John Benjamins.
- Traugott, Elizabeth C. 1988. Pragmatic strengthening and grammaticalization. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 14: 406-416.
- Traugott, Elizabeth C. 1989. On the rise of epistemic meanings in English: an example of subjectification in semantic change. *Language* 65: 31-55.
- Traugott, Elizabeth C. 1995. Subjectification and grammaticalization. In Dieter Stein and Susan Wright (eds.), *Subjectivity and Subjectification: Linguistic Perspectives*, 31-51. Cambridge: Cambridge University.

- Traugott, Elizabeth C. 2001. Legitimate counterexamples to unidirectionality. Paper presented at Freiberg University, October.
- Traugott, Elizabeth C. 2003. Constructions in grammaticalization. In B. Joseph and R. Janda (eds.), *A Handbook of Historical Linguistics*, 624-47. Oxford: Blackwell.
- Traugott, Elizabeth C., and Bernd Heine, eds. 1991. *Approaches to Grammaticalization*. Amsterdam: Benjamins, 2 vols.
- Traugott, Elizabeth C., and Richard Dasher. 2004. *Regularity in Semantic Change*. Cambridge: Cambridge University.
- Trudgill, Peter, and J. Chambers. 1991. *Dialects of English: Studies in Grammatical Variation*. London: Longman.
- Vaux, Bert. 1999/2005. *Harvard Survey of North American Dialects*. Cambridge: Harvard University.
- Vaux, Bert, and Marius L. Jøhndal. 2009. Cambridge Online Survey of World Englishes. < <http://www.ling.cam.ac.uk/survey/>>. Cambridge: Cambridge University.
- Waksler, Rachelle. 1995. She's a mensch and he's a bitch: neutralizing gender in the 1990s. *English Today* 11: 3-6.
- Wales, Katie. 1983. *Thou* and *You* in Early Modern English: Brown and Gilman reappraised, *Studia Linguistica* 37.2: 107-25.
- Wales, Katie. 1996. *Personal Pronouns in Present-Day English*. Cambridge: Cambridge University Press.
- Wales, Katie. 2004. Second person pronouns in contemporary English. *Franco-British Studies* 33.4: 172-85.

Wales, Katie. To appear. Second person pronouns in contemporary English: end of a story or just the beginning? University of Leeds. Unpublished ms.

Walker, Terry. 2003. *You and thou* in Early Modern English dialogues: patterns of usage. In Irma Taavitsainen and Andreas Jucker (eds.), *Diachronic Perspectives on Address Term Systems*, 309-342. Amsterdam: John Benjamins.

Wray, Allison, and Revere Perkins. 2000. The functions of formulaic language: an integrated model. *Language and Communication* 20: 1-28.

Corpora

Davies, Mark. 2007+. Time Magazine *Corpus* (TIME): 100 million words, 1920s-2000s. Available online at <<http://corpus.byu.edu/time>>. [used for examples only]

Davies, Mark. 2008+. *The Corpus of Contemporary American English* (COCA): 425 million words, 1990-present. Available online at <<http://www.americancorpus.org>>.

Davies, Mark. 2010+. *The Corpus of Historical American English* (COHA): 400+ million words, 1810-2009. Available online at <<http://corpus.byu.edu/coha>>.

Kroch, Anthony, and Ann Taylor. 2000. *Penn-Helsinki Parsed Corpus of Middle English*, second edition. <<http://www.ling.upenn.edu/hist-corpora/PPCME2-RELEASE-3/index.html>>.

Kroch, Anthony, Beatrice Santorini, and Ariel Diertani. 2004. *Penn-Helsinki Parsed Corpus of Early Modern English*. <<http://www.ling.upenn.edu/hist-corpora/PPCEME-RELEASE-2/index.html>>.