PRAGMATICS AND THE SOCIAL LIFE OF THE ENGLISH DEFINITE ARTICLE

ERIC K. ACTON

Eastern Michigan University

Eckert (2008) rightly points out that context, variation, and indexicality are inextricably bound. This work—an in-depth case study of the social significance of the English definite article—presents a picture whereby semantic meaning is part of that same web of interrelations. The primary empirical claim of this work is that using *the* with a plural NP (e.g. *the Americans*) to talk about all or typical members of a group of individuals tends to depict that group as a monolith separate from the speaker, and to an extent that using a bare plural (e.g. *Americans*) does not. I present two variationist, corpus-based studies that provide clear evidence of this effect. I then provide a principled account of the effect, building on the insights of sociolinguistic and pragmatic research and extending their collective reach. As I show, the effect is largely rooted in crucial differences between the semantic meaning of *the*-plurals and that of related alternative expressions. As with a broad range of associated phenomena, the exact interpretation of a particular *the*-plural on a given occasion of use depends importantly upon its indexical character, the beliefs of the speech participants, and myriad other contextual factors, but is nonetheless constrained in a principled way.*

Keywords: pragmatics, sociolinguistics, definiteness, social meaning, corpus linguistics, semantics

'Let me be clear: I don't trust the Republicans … '
—US Republican Senator Ted Cruz, 2013

1. INTRODUCTION. The English definite article is most assuredly not a word to be trifled with. Research on the semantics of definites predates the invention of the talking picture, and to date there is a lack of consensus even as to whether *the* carries an existence presupposition (Coppock & Beaver 2012), let alone agreement concerning uniqueness and familiarity (e.g. Strawson 1950, Birner & Ward 1994, Roberts 2002, Abbott 2008, Elbourne 2013).

The demands the respect not only of the linguist and the grammarian but of the everyday English speaker as well. Indeed, speakers must wield *the* with care, for the use of this unassuming function word, commonest of all English expressions, can in fact send potent social signals. Reference to one's own spouse as 'the wife', for example, paints a different picture of the speaker's marriage than does a reference to 'my wife'. At the very least, the former suggests extra distance between speaker and referent relative to the latter. And it is this same basic dynamic that US Republican Senator Ted Cruz exploited in his ironic remark in the epigraph above, suggesting an outsider status vis-à-vis his own party. Example 1a likewise exhibits this meaning, where *the Americans* generally suggests that the speaker of 1a is not an American. Absent *the*, as in 1b, the effect is diminished.

(1) a. **The Americans** love cars.
    b. **Americans** love cars.

This work—part of a larger project relating semantics, pragmatics, and sociolinguistics—presents an in-depth study of the social meaning of *the*. Specifically, I show that using *the* with a plural NP (a '*the*-plural') to talk about all or typical members of a group

of individuals tends to depict that group as a monolith of which the speaker is not a member—and to an extent that using a bare plural (BP) does not.

Although using a *the*-plural very often does suggest that the speaker is not a member of the relevant group, is deemphasizing their membership in the group, or is emphasizing their nonmembership[1]—what I call the DISTANCING EFFECT of *the*-plurals—the meaning is not an entailment. In 2, for instance, it is clear that *the Tea Party Patriots* and the possessors implicit in *our* pertain to the same group of individuals.

    (2) **The Tea Party Patriots** depend on grassroots activism where it counts: near you. Local Groups are the backbone of **our** activities across the nation.
                          (from the website of the 'Tea Party Patriots')[2]

Being nonentailed, the meaning of interest does not fit squarely within the purview of semantics. This suggests a need to turn to the subdisciplines that address nonentailed meaning: namely pragmatics and, increasingly, sociolinguistics—particularly sociolinguistics in the THIRD WAVE of variation studies (Eckert 2012). Indeed, doing justice to the present phenomenon—and a broad class of related phenomena besides—means drawing from, connecting, and extending both traditions.

Concerning sociolinguistics, the meaning of interest is social in nature in that it concerns the speaker's relation to other individuals. Moreover, as I show in §3, the phenomenon is amenable to variationist research. Accordingly, sociolinguistics has much to offer here. However, the present case departs from the vast majority of variationist sociolinguistic research, for which the relevant variants do not bear propositional meaning themselves, or for which it is assumed 'that distinctions in referential value among [variants] are neutralized in discourse' (Tagliamonte & Denis 2014:97, citing Sankoff 1988:153; see also Romaine 1984, Cheshire 2005). Instead, the meaning of interest crucially depends upon differences between the encoded, semantic meaning of *the*-plurals and that of related alternative forms.

Accordingly, pragmatics too has much to offer in the present case. Pragmatic research has greatly illuminated the dynamics by which utterances are enriched with nonentailed meanings (e.g. Grice 1975, Horn 1984, Sperber & Wilson 2004). At the same time, the present phenomenon (and many others like it) calls for extending the insights of previous research and complementing them with sociolinguistic theory and methodology.

The remainder of this article is organized as follows. I first provide a more thorough overview of the meanings of interest (§2), and then present two corpus-based, variationist studies that show that the effect is clearly reflected in usage patterns and is sensitive to contextual factors (§3). In §4, I provide an overview of the insights that the traditions of pragmatics and sociolinguistics offer for explaining the distancing and monolithizing effects of *the*, and connect and extend these insights in §5 and §6 in order to explain these effects. As I show, a full account calls for an analysis that is at once semantic, pragmatic, and sociolinguistic. Section 7 takes stock of the findings and the analytical approach of the present article, looking toward future sociolinguistic, pragmatic, and semantic research.

2. MORE EXAMPLES OF THE MEANING. As discussed with respect to example 1 and the epigraph from Ted Cruz, the use of *the* in talking about individuals often engenders a distancing effect. A 1958 book of photographs of life in the US titled *The Americans* provides another sharp illustration. Authored by Swiss-born photographer Robert

---

[1] I thank a referee for the suggestion to foreground the set-membership basis of the distancing effect.
[2] 'Tea Party Patriots'; https://www.teapartypatriots.org/, last accessed July 10, 2015.

Frank, who immigrated to the US in 1947, '[i]t was initially dismissed as the jaundiced work of an unpatriotic cynic' (McAlley 2009:n.p.). Arts critic Richard B. Woodward writes: 'What was most upsetting about Mr. Frank's take … was, first of all, the provocative title. It is as though the people in his pages were an alien species and he a more evolved anthropologist' (Woodward 2009:n.p.). The title of a 2009 exhibition of Frank's work at the National Gallery of Art, 'Looking in: Robert Frank's "The Americans" ', further depicts Frank as an outsider.

The same dynamics underlie the remarkability of Republican Donald Trump's reportedly frequent use of the phrase *the Republicans*. An article from the *Economist*, aptly titled 'Outsiders' chance', observes: 'Donald Trump … was once a registered Democrat and still refers derisively to his party as "the Republicans", as if it is some unpromising acquisition he has been arm-twisted into buying' (Outsiders' chance 2016).

Additional evidence of the distancing effect of *the* can be found in its use in disparaging generalizations, discourse on marginalization, descriptions of social oppositions, and 'us-versus-them' rhetoric.

(3) Faced with the prospect of a war against all, humans reformulate the conflict into a war of nearly all against a few—against **the Jews** or **the communists** or **the gays** or **the feminists** or **the Mexican immigrants**.   (Placher 2009)

(4) Paid for by Bernie 2016 (not **the billionaires**).[3]
          (from US Senator Bernie Sanders's presidential campaign website)

(5) And I just ask some of you that are back on the East Coast, come down to the border and talk to those of us that have seen what happens when **the politicians** play politics with something that ends up costing people's lives …
          (Republicans backed into a corner 2012)

In each example, the relevant groups are depicted as separate, removed, or opposed. Moreover, I submit that in every case the removal of *the* would dull the tone of separation or derogation to some extent.

The distancing effect is so pronounced in some cases that certain *the*-plurals seem to have attained taboo or near-taboo status. In discourse surrounding race and ethnicity in the US, *the blacks* is a clear example. Indeed, the phrase *the blacks* is virtually absent in contemporary mainstream US news media, despite myriad tokens of *blacks*. In the eleven months following the killing of Michael Brown in Ferguson, Missouri, for instance, whereas in the *New York Times* and the *Washington Post* tokens of the BP *blacks* in articles containing the word *Ferguson* number in the hundreds, there are only two such tokens of *the blacks*. Tellingly, both tokens occur not in the authors' own words but in quotations, and one of them, provided as example 6, is especially transparent in its prejudice.

(6) 'They always want to stir up to trouble, the blacks.'   (Robertson 2014)

A similar pattern holds for *the gays*, for which the effect is strong enough as to be employed in representations of individuals presumed to be unsympathetic toward or unfamiliar with gay people. Example 7, from the review of a play on gay politics in 1980s Great Britain, illustrates.

(7) When a curious older woman insists on calling Mark and his friends '**the gays**' and asks wide-eyed questions about all gay women being vegetarians, the naïveté could be dismissed as a cheap and easy laugh.   (Hornaday 2014)

---

[3] https://berniesanders.com/, last accessed June 30, 2016.

Of course, in many cases the effect is not so pronounced and problematic. In fact, the distancing meaning can be canceled or overshadowed by other considerations (as in example 2), a matter about which I have more to say later. Generally, however, using a *the*-plural does tend to have a distancing effect (as in examples 1 and 3–7) and, I claim, to a greater extent than the use of a BP. If this is so, it should be reflected in usage patterns. For one thing, we ought to find that, ceteris paribus, speakers' *the*-plural-to-BP ratio should be higher when talking about groups of which they are nonmembers or from which they wish to express distance than when talking about groups of which they are a part or for which they have an affinity. In the next section, I show via two corpus-based variationist studies that this is indeed the case.

3. TWO VARIATIONIST STUDIES. I begin with a study of the use of *the*-plurals and BPs by members of the US House of Representatives to refer to Democrats and Republicans over a twenty-year period. In accordance with the discussion above, representatives have a much higher *the*-plural-to-BP ratio in talking about their opposing party than in talking about their own. I then present an analysis of terms used on the political talk show *The McLaughlin Group* to talk about Democrats and Republicans that provides further evidence of this pattern. At the same time, key differences between the two corpora lead to results that differ in other, predictable ways, underscoring the importance of context in the use and significance of these expressions.

By linking patterns of variation to differences in the social significance of the variants of interest (in this case *the*-plurals and BPs), this work bears the defining trait of what Eckert calls the 'third wave' of sociolinguistic variationist research (see §4.2). At the same time, the present work departs from the mainstream of variationist research in that the variants of interest are not equivalent even in their propositional meanings. They can, however, be used for similar communicative purposes—namely for talking about all or typical individuals of a particular type. It is under such circumstances that the two kinds of expression may be reasonably treated as alternatives to one another (Romaine 1984, Walker 2010, Cameron & Schwenter 2013), thereby serving as an appropriate basis for variationist study. But the underlying semantic differences between these expression types must not be glossed over, despite their shared communicative function under the relevant circumstances. On the contrary, I later show that these differences are the very kernel of the social meaning of interest.

3.1. *THE* IN THE US HOUSE OF REPRESENTATIVES. To provide empirical evidence of the distancing effect of *the*, I turn to Djalali's (2013) HOUSE PROCEEDINGS CORPUS (HPC), a complete set of transcripts of the proceedings of the US House of Representatives from February 1993 through December 2012. The HPC provides an ideal testing ground for the hypothesis, given its size (tens of millions of words), its diversity of participants (over 800 distinct speakers), and the partisan nature of the discourse.

The principle measure of interest is in 8, henceforth the *the*-percentage or *the*-%.

(8) For any plural count noun *Xs* and set of speakers *S*, the THE-PERCENTAGE (*THE*-%) for *Xs* for *S* is defined as:
$\Sigma_{s \in S}$ (# of s's tokens of *the Xs*) / $\Sigma_{s \in S}$ (# of s's tokens of *the Xs* + # of s's tokens of BP *Xs*)

Based on this measure, the hypothesis for this section is given in 9.

(9) HYPOTHESIS: On average, representatives will have a higher *the*-% when talking about the opposing party than when talking about their own.

CATEGORIZING TOKENS FOR INCLUSION. There are more than 70,000 tokens of *Democrats* and *Republicans* in the HPC, so analyzing each token by hand is time-prohibitive. The following methodology was adopted instead.

The initial data set included every sentence in the HPC containing at least one token of *Democrats* or *Republicans*. The corpus contains a handful of duplicate transcripts; any sentences that were duplicative in terms of the sentence itself, the speaker name, and the date of utterance were removed. Among the remaining sentences are 73,242 tokens of *Democrats* and *Republicans*. To simplify the analysis slightly, only the first token of *Democrats* in sentences containing *Democrats* was analyzed, and similarly for *Republicans*, so that 71,199 (97.2%) of all tokens were included in the analysis.

The next step was to identify the determiner for each token. Tokens of *Democrats* or *Republicans* for which the noun was pre- or postnominally modified (e.g. *Senate Democrats*, *Republicans in the House*) were excluded for two reasons. First, restricting the analysis to unmodified tokens helps ensure that the tokens are comparable; there is no doubt a major difference between, say, *(the) blue-collar Democrats* and *(the) Democrats*. Second, admitting modified tokens would allow for cases in which the speaker is not a member of the subgroup being talked about, but is a member of the larger group associated with the head noun—complicating the picture significantly. The determiners, quantifiers, and partitives coded for are enumerated in 10.

> (10) all Arabic numerals 2 through 999 (with or without *of the*), *zero*, all written numeral terms *two* through *twenty*, and the following expressions:
> any (of the), all (of the), enough (of the), few (of the), fewer (of the), many (of the), most (of the), more (of the), no, none of the, other (of the), several (of the), some (of the), the, these, those, my, our, your, his, her, its, their, us, we, you

Any token not immediately preceded by one of the items in 10 was coded as bare. Some such tokens had a determiner with a modifier intervening, but, as stated above, modified tokens were ultimately excluded from the analysis. Tokens preceded by the modifiers in 11 were excluded, as were tokens followed by the items in 12. Each item in the lists was included on the basis of an actual token of *Democrats* or *Republicans* preceded/followed by that item in the HPC.

> (11) committee, subcommittee, House, Washington, senate, party, well-respected, Congressional, Texas, California, southern, regular, sensible, senior, junior, rank-and-file, moderate, conservative, liberal, centrist, right-wing, left-wing, radical, extremist, fat-cat, blue(-)collar, blue(-)dog, -and-spend, brave, new, judiciary, fellow, other(-)body, good, bad, teenage, Christian, thoughtful, mean-spirited, valiant, responsible, -life, -choice, key, Bush, Clinton, Reagan, Gingrich, majority, minority, ruling, ranking, white, black, middle class, poor, wealthy, leading, young, freshm(e/a)n, do-nothing, 30-something

> (12) who, on the, on this, on that, on these, on those, in the, in this, in that, in these, in those, in Congress, in Washington, that are, that were, that will, that said, that voted, that promised, up here

I further excluded tokens for which the relevant DP was the pivot of an existential *there* as in 'there are Democrats on the panel', because such tokens generally cannot be substituted with *the*-plurals, and where they can, a substantial change in meaning results. Finally, any tokens for which the speaker was unidentified or neither a Democrat nor a Republican were excluded. The remaining 54,393 tokens—30,787 BPs and 23,606 *the*-plurals—served as the basis for the *the*-% calculations.

To test the accuracy of this process, a random sample of 250 of the original 71,199 sentences was checked by hand. The automated procedure performed very well in terms of both precision and recall. Among the 250 sample tokens, 213 were marked for inclusion via the automated procedure. Of those 213, 96.7% were both correctly marked for inclusion and assigned the correct determiner. There were an additional three tokens

marked for exclusion that ought to have been marked for inclusion, so that, in all, 206 (98.1%) of the 210 true positives were retrieved and correctly categorized.

The *the*-%s based on the manual and automated labeling methods differed minimally. For every combination of speaker party and party term, the ratio of the *the*-%s calculated based on the two different methods was between 0.96 and 1.02. The only case in which the results based on automated labeling were more favorable to the present hypothesis was Republicans' *the*-% for the term *Democrats*, where the manual approach yielded 45.3%, and the automated labeling yielded 47.2%. As the results below show, this potential bias has no substantial bearing on the evaluation of the hypothesis.

AGGREGATE-LEVEL FINDINGS. The results overwhelmingly support the hypothesis in 9. Table 1 presents the aggregate *the*-%s for both parties for both party terms, based on over 10,000 tokens for each combination of speaker party and party term. Democrats' *the*-% for the term *Democrats* is 30.4% compared with 54.4% for the term *Republicans*. Conversely, Republicans' *the*-% for the term *Democrats* is 53.3%, compared with 26.1% for the term *Republicans*. This means that each party's *the*-% for the opposing party term is over 1.75 times higher than the *the*-% for their own party term. This holds even if one scales down the *the*-% of *Democrats* for Republicans by 4.0% to 51.2% in order to account for the potential bias introduced by the automated labeling (see above).

SPEAKER PARTY	DEM *the*-%	REP *the*-%	DEM N	REP N
Democratic	30.4%	54.4%	11,352	18,992
Republican	53.3%	26.1%	13,007	11,042

TABLE 1. Aggregate *the*-%s for US House representatives by party.

MODELING THE DATA. A pair of exploratory mixed-effects models further confirm that speaker party has a significant and dramatic effect on the use of *the*-plurals/BPs in the HPC. A comprehensive investigation of what other factors condition the use of these expressions would make for a productive study in its own right. In the meantime, I present the following initial investigation.

As a step toward uncovering other contributory factors, the 250 tokens used to validate the coding methodology were examined for additional distributional patterns. Two clear trends emerged. First, coordination of nominal phrases containing *Republicans* and *Democrats* with *and* highly favors BPs over *the*-plurals in the sample—only two of the thirty-five relevant instances of coordination involved *the*. In addition, sentences in the sample that include cooperative language also favor BPs over *the*-plurals. In the eighteen relevant sentences containing *together*, *alike*, *bipartisan(ship)*, or *work(s/ed/ing) with*, every one of the DPs of interest was a BP.

To test the robustness of these observations in the larger corpus, I used generalized linear mixed-effects modeling to predict whether a given token of *Democrats* occurred as a BP or as part of a *the*-plural on the basis of speaker party, whether the token was conjoined with *Republicans* via *and*,[4] and whether the carrier sentence contained *together*, *alike*, *bipartisan(ship)*, or *work(s/ed/ing) with* (an attempt to code for cooperative language based on the operative words in the random sample).

Speaker identity was included in the model as a random-effect group (Johnson 2009, Tagliamonte & Denis 2014). Given the exploratory nature of this analysis, nested mod-

---

[4] In cases of coordination, if *the* preceded only the first conjunct (less than 3% of all relevant cases of coordination in the data set), as in *the Republicans and Democrats*, it was treated as taking scope over both.

els including each possible combination of (i) the interaction terms provided for by the three independent variables and (ii) random-slope terms based on presence of coordination or cooperative language were compared for statistical performance.[5] Models were run using the glmer() function in R (Bates et al. 2015, R Core Team 2012) and compared using R's anova() function. The reported model is the simplest model that did not perform significantly worse than any other more complicated model ($\alpha = 0.05$). The results are provided in Table 2 under 'model 1', where positive coefficient estimates (column 'est.') favor the use of a *the*-plural over a BP.

The estimates for 'Speaker party-R', '*And*-Y', and 'Cooperative-Y' correspond to the effects of the speaker being a Republican, coordination with the other party term via *and*, and cooperative language in the carrier sentence, respectively. 'Party-R : *And*-Y' corresponds to the additional effect of coordination with *and* in Republicans' talk about Democrats.

	MODEL 1: *(the) Democrats* (*N*: 24,359; speakers: 833)				MODEL 2: *(the) Republicans* (*N*: 30,034; speakers: 829)			
FIXED EFFECTS	EST	SE	z	Pr(>|z|)	EST	SE	z	Pr(>|z|)
(intercept)	−0.707	0.061	−11.59	< 0.001	0.341	0.060	5.67	< 0.001
Speaker party-R	1.268	0.084	15.13	< 0.001	−1.173	0.087	−13.40	< 0.001
And-Y	−2.016	0.114	−17.72	< 0.001	−2.992	0.108	−27.68	< 0.001
Cooperative-Y	−0.829	0.070	−11.82	< 0.001	−0.803	0.088	−9.10	< 0.001
Party-R : *And*-Y	−0.820	0.153	−5.36	< 0.001	1.374	0.151	9.11	< 0.001
Party-R : Coop-Y	—	—	—	—	0.273	0.134	2.04	0.041
RANDOM EFFECTS								
Speaker (intercept)	Var.: 0.693	SD: 0.833			Var.: 0.870	SD: 0.933		
And-Y	Var.: 0.442	SD: 0.665			Var.: 0.389	SD: 0.624		
LIKELIHOOD/DEVIANCE	logLik: −13,045				logLik: −15,974			
	AIC: 26,106				AIC: 31,948			

TABLE 2. Models predicting *the*-plural (rather than BP) in the HPC.

All effects are significant ($\alpha = 0.001$) in the expected direction. Considering tokens of *Democrats* for which the token is neither coordinated with a *Republicans*-DP nor attended by cooperative language (75.9% of all relevant tokens), the model predicts a massive effect for speaker party, whereby the *the*-% for Democrats is 33.0%, compared with 63.7% for Republicans. At the same time, coordination of the relevant kind so greatly favors BPs as to nearly overwhelm the speaker-party effect where it does occur—bringing the predicted rates down to 6.2% and 9.3%, respectively, where there is no attendant cooperative language, and 2.9% and 4.3% where there is.[6]

The results of an analogous model designed to predict the distribution of *Republicans* and *the Republicans*, selected by the same means as the first, is presented in Table 2 under 'model 2'. This model is nearly identical to model 1 in structure, the only difference being that here the interaction between speaker party and cooperative language just makes the

[5] Random-slope terms based on speaker party were excluded, as only four of the more than 800 speakers in the corpus changed their political party during the relevant time frame, none of whom had enough tokens to draw any substantial conclusions about their *the*-%s before and after changing parties. For simplicity, random slopes based on interactions were not pursued. lme4 package version 0.999999-0 was used.

[6] As for random effects, in this model and the next the standard deviation values for each random effect relative to its corresponding fixed-effect estimate indicates a fair amount of interspeaker variation in terms of baseline rates of *the*-plural use and, to a lesser extent, the effect of coordination.

cut for statistical significance ($p = 0.041$), such that the presence of cooperative language lowers Democrats' *the*-% for *Republicans* more than it lowers Republicans'.[7]

As in model 1, the effect of speaker party is enormous in cases where the relevant token is neither conjoined in the relevant way nor attended by cooperative language (80.1% of cases), so that the predicted rate of *the Republicans* is 58.5% for Democrats and 30.3% for Republicans. Also as before, coordination of the relevant kind so strongly favors BPs as to effectively overwhelm the effect of speaker party in cases where it is present—especially when working in concert with cooperative language, where the *the*-%s predicted by the model are 3.1% and 4.8%, respectively.

Fully accounting for the effects of cooperative language and coordination on the use of plural party terms will require further research. In the meantime, I submit that even these effects may in large part be explained by the same basic difference in social meaning underlying the speaker-party effect. Inasmuch as *the*-plurals, relative to BPs, depict groups of individuals as distant, separate monoliths, it stands to reason that BPs are better suited for messages citing or encouraging cooperation or togetherness between such groups, ceteris paribus. This, as the models suggest, is precisely what we find in the HPC, as in example 13.

(13) These are problems that **Republicans** are anxious to **work with Democrats**
      on.   (Dana Rohrabacher (R-CA), 27 October 2009)

As for sentences in which DPs containing the two party terms are conjoined via *and*, such sentences typically suggest a commingling of both parties—similarly disfavoring distancing, boundary-drawing *the*-plurals. Indeed, such sentences often contain cooperative language themselves, thus doubly conveying an inclusive, boundary-less tone, as in 14.

(14) **We all agree**—all of us, **Republicans and Democrats alike**—that cuts in
      wasteful spending are vital to our country's future.
                                            (William Keating (D-MA), 1 March 2011)

Furthermore, the fact that coordination of party terms in this corpus entails naming one's own party may further favor BPs in such sentences: using a *the*-plural for one's opposing party, together with syntactic like-conjoins-with-like constraints (e.g. Levy 2015), would pressure speakers to use a *the*-plural for their own party, which is generally undesirable for reasons again tied to the same social meaning. Of course, other factors like considerations of rhythm and concision may well have a role to play, too. In any case, it seems that even the effects of the presence of coordination and cooperative language are bound up with the same contrast in social meaning between BPs and *the*-plurals as is the effect of speaker party.

DISCUSSION. The foregoing analysis presents a dramatic pattern whereby members of both parties use *the*-plurals rather than BPs at far higher rates when talking about the opposing party than when talking about their own. This is exactly what we would expect under the central empirical claim of this work—that is, that *the*-plurals, more so than BPs, tend to depict the group of individuals being talked about as a monolith of which the speaker is not a part. The significant effects of cooperative language and coordination with *and* provide even further evidence of the monolithizing, distancing ef-

---

[7] This is not especially surprising. As with coordination with *and*, insofar as cooperative language brings *the*-%s down to low levels for both parties, the effect would have to be larger for Democrats, whose baseline *the*-% for *Republicans* is much higher than Republicans'. The analogous effect is marginally significant in model 1 ($p = 0.095$).

fect of *the*-plurals. We thus have a clear case of principled variation, whereby differences in meaning beget differences in use along social lines.

On exceptions. Despite the large effect size of speaker party, it is worth noting that some individuals in the HPC have *the*-%s that run against this general pattern. Among the representatives in the corpus who had at least ten tokens of both *(the) Democrats* and *(the) Republicans*, 12.4% of Democrats and 13.8% of Republicans had at least as great a *the*-% for their own party as for their opposing party. These numbers are relatively small, but not negligible.

Of course, one's party is not the only factor influencing one's choice of DP. Some of these speakers, for instance, used *the*-plurals in indirect quotation, in contexts where a *the*-plural might be expected because the quoted individual is not a member of the relevant party.[8]

(15) People will say we could solve the problem of deficits if only **the Democrats** or the Congress would hold down spending.

<div align="right">(Brian Baird (D-WA), 26 February 2003)</div>

(16) I want to respond to the comments [… from] some of my colleagues from across the aisle, in which they have said **the Republicans** did not participate in passing social security.   <div align="right">(Vernon Ehlers (R-MI), 11 August 1994)</div>

Other tokens were direct quotations, where the speaker had no choice as to what DP to use (assuming faithfulness to the original source). And still others were issued ironically, such that the speaker mocked the perspective of someone from the opposing party.

(17) Anyone who objects to doing for Europe what European boys should be doing naturally despises children almost as much as **the Republicans** hate old folks …   <div align="right">(Robert Dornan (R-CA), 28 November 1995)</div>

Moreover, the monolithizing effect of *the*-plurals lends them well to cases in which one wishes to present one's party as a solid, unified front, as in 18 and 19.

(18) **The Democrats** are united on the need for a new direction in Iraq.

<div align="right">(John Larson (D-CT), 26 June 2006)</div>

(19) What **the Republicans** are going to stand against is tying the funds our soldiers need to do their jobs to benchmarks thought up by special interest groups.

<div align="right">(Adrian Smith (R-NE), 9 March 2007)</div>

And, as observed above, coordination of the two party terms and the presence of cooperative language favor BPs as well.

Speakers whose utterances about their own party were disproportionally influenced by any of these or other relevant considerations may well have had a higher *the*-% for their own party even if they strongly identified with their party. Moreover, the picture is further complicated for any speakers who have a complex relationship with their party.

Further investigation of the speech of the minority of representatives who spoke in unexpected ways must await future research. The important points for the present are that the vast majority of speakers in the HPC patterned as predicted, and that multiple factors (both linguistic and social) play a role in the use of *the*-plurals and BPs.

---

[8] A referee observes that such cases suggest that the distancing meaning can be embedded. Potts (2007) notes similar cases of embedding with expressives like *bastard* and accounts for them by endowing expressives with a context-sensitive judge parameter. That analysis might not apply as well to the present case—whereas expressing some judge's emotional state or evaluative stance is the raison d'être for expressives (thus calling for a judge parameter), this is not so for *the*-plurals. Due to limited space, however, I must leave further exploration of the theoretical consequences of embedded-but-nonentailed social meanings for future research.

The thoroughly partisan nature of the HPC makes it ideal for seeing the social meaning of interest in action. But there are also lessons to be learned from investigating the same issues in a corpus that differs from the HPC in important ways. For that, I turn now to a study of talk about Democrats and Republicans on the talk show *The McLaughlin Group*.

3.2. *The* ON THE MCLAUGHLIN GROUP. *The McLaughlin Group* (aired 1982–2016) was a weekly political talk show featuring the discussion of topical issues among five pundits, including host John McLaughlin, who moderated and added commentary of his own. The corpus for this study (hereafter, the MGC) consists of the 154 consecutive episodes airing from May 23, 1998, through May 13, 2001, and contains over 700,000 words. As I show below, the distancing effect of *the* obtains in this corpus as it did in the HPC. Due to crucial differences between the two corpora, however, certain aspects of the results are importantly different.

KEY DIFFERENCES AND PREDICTIONS. The first key difference is that whereas the HPC consists of speech from individuals transparently speaking as representatives of their respective parties, the MGC, being based in a journalistic program, involves greater pressure on participants to exhibit objectivity and avoid overt displays of partisanship.

The second, related difference is that the pundits in the MGC, in talking about Democrats and Republicans, are often talking about Democratic and Republican politicians in particular, and, unlike the speakers in the HPC, none of the pundits in this study served as an elected official during the relevant time span. Thus, wherever references to '(the) Republicans' or '(the) Democrats' are meant to say something about Republican or Democratic politicians specifically, MGC participants are speaking as outside observers regardless of their party affiliations, whereas HPC participants are often talking about groups of which they truly are members.

These two differences yield two predictions. First, given the pressure for MGC speakers to temper their political biases, we should expect that the difference between a given speaker's *the*-%s for the two parties (*THE*-% DIFFERENTIAL) will generally be smaller in the MGC than in the HPC, as a large differential could signal heavy bias in favor of one party over the other. Second, since in referring to '(the) Democrats' and '(the) Republicans' MGC speakers more often inhabit the role of the outside observer than HPC speakers do, we should expect that MGC speakers' *the*-%s will generally be as high as or higher than those of HPC speakers.

METHODS. This study centered on the eight pundits (including McLaughlin) who appeared on at least twenty-five of the MGC's 154 episodes. Participants' political leanings were determined based on their careers, self-descriptions, and political statements and categorized based on the basic, if imperfect, conservative ≈ Republican and liberal ≈ Democrat dichotomy (Saad 2015).[9]

For each participant, every token of *(the) Democrats* and *(the) Republicans* was collected, excluding tokens for which the head noun was modified, for the sake of comparability. Tokens were identified by hand. I retained tokens for which the DP of interest

---

[9] Here is a snapshot of each speaker's political profile. Michael Barone: American Enterprise Institute fellow, *Washington Examiner* analyst. Tony Blankley: Press secretary to Newt Gingrich (1990s). Pat Buchanan: Sought Republican presidential nomination 1992, 1996. Larry Kudlow: Author of *Bullish on Bush* (Kudlow 2014). John McLaughlin: Part of Nixon and Ford administrations (Maltin 2005:n.p.). Eleanor Clift: Bill and Hillary Clinton supporter (Clift 2008:n.p.). Lawrence O'Donnell: Self-described socialist, but 'as much a capitalist as a socialist' (O'Donnell 2011:n.p.). Clarence Page: '[F]actory-town liberal' (Page 2014:n.p.).

was the object of the word *of*, provided that either variant would be grammatical in the context (e.g. *a lot of (the) Democrats*). Excluding such tokens does not change the results in any appreciable way. Quotative uses were excluded, and *the*-%s were calculated for each speaker.

RESULTS. As shown in Table 3, the use of *the*-plurals patterns according to one's closer political party in seven of the eight speakers, and on the aggregate by party.

PARTICIPANT	CLOSER PARTY	DEM *the*-%	REP *the*-%	DEM *N*	REP *N*	Lower *the*-% for closer party
Barone	R	67.4%	50.0%	46	36	yes
Blankley	R	82.7%	53.6%	75	84	yes
Buchanan	R	71.4%	66.7%	14	36	yes
Kudlow	R	62.5%	56.3%	16	16	yes
McLaughlin	R	52.8%	45.6%	199	204	yes
Clift	D	60.4%	62.2%	106	148	yes
O'Donnell	D	47.4%	45.5%	19	22	no
Page	D	28.6%	66.7%	14	15	yes
Closer to Republican		62.3%	50.3%	350	376	yes
Closer to Democratic		55.4%	60.5%	139	185	yes

TABLE 3. *The*-%s for MGC participants, ordered by political leaning and surname.

In addition, as expected given MGC pundits' role as outside observers, *the*-%s are higher on average in this corpus than in the HPC, as shown in Table 4. Aggregating by closer party and party term, we see that the *the*-% is at least six percentage points higher in the MGC than in the HPC across the board. The effect is especially clear when speakers talk about their own party—in both cases the average *the*-% for the party closer to the speaker is more than 24 percentage points higher in the MGC.

CLOSER PARTY	CORPUS	DEM *the*-%	REP *the*-%	DIFFERENTIAL (abs. value)
Republican	HPC	53.3%	26.1%	27.2%
	MGC	62.3%	50.3%	12.0%
	difference	*−9.0%*	*−24.2%*	*15.2%*
Democratic	HPC	30.4%	54.4%	24.0%
	MGC	55.4%	60.5%	5.1%
	difference	*−25.0%*	*−6.1%*	*18.9%*

TABLE 4. Comparison of aggregate *the*-%s in the HPC and MGC.

Also as expected given pressures for MGC pundits to display a degree of objectivity and nonpartisanship, the aggregate *the*-% differential for a given party is far narrower in the MGC than in the HPC. In fact, as shown in Table 4, this differential in the HPC is more than double that in the MGC for both parties on the aggregate—27.2% versus 12.0% for Republicans/conservatives, and 24.0% versus 5.1% for Democrats/liberals.

Evidence of expectations of objectivity can be found in the corpus itself, as in 20.

(20) MCLAUGHLIN: … Have the Republicans already out-strategized the Democrats? Tony Blankley?

BLANKLEY: In the interest of honest journalism, I don't think we've out-strategized the Democrats since 1994. However, I think the Republicans are sort of stumbling into a defensible position … So I think while we haven't out-strategized the Democrats, we're in a pretty useful position.

MCLAUGHLIN: Why do you keep saying 'we'? [Laughter.] Are you identifying yourself with Republicans now […]? [… Y]ou are an analytic jour-

> nalist. You're like Pat Buchanan, and Pat would never identify himself as being a Republican—a 'we'.[10]

Despite an initial gesture toward objectivity ('In the interest of honest journalism'), conservative Tony Blankley personalizes his statement in talking about Republicans as 'we'. McLaughlin makes fun of Blankley's usage as falling outside the standards of analytic journalism, pointing to panelist Pat Buchanan as a shining example of appropriate usage. Indeed, McLaughlin seems to be onto something: Blankley's *the*-% differential for the two party terms is the highest of all of the conservative participants at 29.1, whereas Buchanan's, at 4.7, is the lowest (see Table 3).

In light of the clear pressure toward journalistic objectivity, it is not surprising to find that participants' *the*-% differentials in the MGC were on average far narrower than those in the HPC.

3.3. SUMMARY. The findings of these studies clearly converge in support of the notion that *the*-plurals, particularly relative to BPs, have a distancing, monolithizing effect. The vast majority of speakers in both corpora had a lower *the*-% for the party they are closer to than for the other. And, on the aggregate, speakers in the MGC had far higher *the*-%s for both party terms (reflective of their status as outside observers) and narrower *the*-% differentials (reflective of pressures to exhibit objectivity). The negative correlation between *the*-plurals and cooperative, commingling language further suggests that *the*-plurals depict their referents as separate monoliths.

Before I provide an account for these patterns, it is worth emphasizing that the foregoing findings concerning political party terms are an instantiation of what is predicted to be a more general pattern, where, ceteris paribus, being a member of a group or having an affinity for that group makes one more likely to opt for a BP over a *the*-plural in naming that group, and vice versa. This same general prediction leads us to expect, for instance, that pundits in the MGC, all being US citizens, would be less likely to use a *the*-plural in referring to Americans than in referring to people from regions outside of the US. As one striking example, the *the*-% across all pundits in the MGC for the term *Americans* is 10.1% ($n = 99$), whereas for *Russians* it is 93.8% ($n = 81$).

By the same logic, if we expect two distinct groups of individuals to bear a similar relation to some third group, we should likewise expect their *the*-%s in naming that third group to be similar, ceteris paribus. For instance, perhaps there is less of a difference between Democratic and Republican representatives in their orientation to voters in general than in their orientation to Democrats. If so, then—holding all other factors fixed—we should in turn expect the two parties' *the*-%s for the term *voters* in the HPC to be closer than their *the*-%s for the term *Democrats*.[11] By the same token, differences in *the*-%s for the two parties for a particular group term could well indicate interparty differences in orientation to the group itself. In this way, data concerning an individual or group's *the*-%s for a range of plural terms may provide clues for mapping out that individual or group's social positioning and attitudes. At the same time, such data must be interpreted with caution, given the multitude of factors influencing determiner choice.

---

[10] July 10, 1998, episode of *The McLaughlin Group*; first aired July 11, 1998. From transcript posted at http://www.mclaughlin.com/transcript.htm?id=40, last accessed July 1, 2016.

[11] It appears that this prediction is indeed borne out in the HPC. Whereas the *the*-% differential between Democrats and Republicans in the HPC for the term *Democrats* is about 23 percentage points, a preliminary analysis indicates that the corresponding value for the term *voters* is well under 10. More work is needed to ensure that the data upon which the latter number is based include only tokens of *voters* that have no pre- or postnominal modifier.

Pursuing that line of inquiry must be left for future research. In what remains herein, I show that the distancing, monolithizing meanings documented above are not accidental, but can be derived from the semantic properties of *the*-plurals and related expressions considered in the light of general pragmatic principles.

4. A SOCIOPRAGMATIC PERSPECTIVE. The central phenomenon of interest in this work is a kind of nonentailed meaning. In looking to explain the phenomenon, then, it makes sense to seek answers in the subfields of linguistics most deeply concerned with nonentailed meaning: namely pragmatics and sociolinguistics. Both areas have illuminated a great many phenomena, and both will be of much help here. Doing full justice to the meanings of interest, however, means expanding and connecting both lines of research. In this section, I provide a brief overview of what each tradition has to offer to the study of nonentailed meaning, focusing on matters that are particularly relevant to the present case. Then, drawing on and further developing that discussion, I present a sociopragmatic account of the phenomenon of interest—one that (i) in the spirit of pragmatics, ties the core meanings of interest to principles of rational action and the semantics of *the*-plurals and related alternatives (§5); and (ii) in the spirit of sociolinguistics, addresses the role of context, variation, and indexicality (§6).

4.1. A VIEW FROM PRAGMATICS. Perhaps Grice's greatest contribution to the study of meaning is the idea that much of what is conveyed by an utterance is not entailed but rather arises from considering an observed utterance in relation to related alternative utterances and general principles guiding language use. Underlying this perspective are the principles in 21.

> (21) Language users (whether consciously or not):
>> a. Have context-sensitive expectations about what makes for a normal or appropriate utterance (e.g. Grice 1975, Brown & Levinson 1987, Keller 1994, Traugott 2011).
>> b. Attempt to optimize the cost-benefit ratio of their utterances, ceteris paribus (e.g. Horn 2004, Sperber & Wilson 2004, Goodman & Frank 2016).
>> c. Appeal to reason and take 21a,b into account in forming and interpreting utterances (e.g. Grice 1975, Horn 1984).

The degree to which language users consider or enact these principles consciously is certainly open to investigation. At a minimum, however, the idea is that language users' behavior tends to square with these principles. (Of course, when such principles are not in force, familiar pragmatic phenomena like implicature cease (Horn 2004:7–8).) It is but a short step from these principles to the following very general dynamic, which in some form or another (though not necessarily in the following terms) lies at the foundation of much pragmatic research on nonentailed meaning—including the present study.

Given an utterance u, whenever one may assume that 21b is in force, it follows that there is no alternative utterance α such that the speaker believed that α would offer a better mix of costs and benefits than u in the context of utterance. Now if from the perspective of an observer of u there is some α' that, on the face of things, appears that it might offer an appreciably better mix of costs and benefits than u for the speaker, this suggests room for improvement in the observer's model of the speaker's beliefs, goals, and so on—otherwise it would be plainly obvious why the speaker preferred u to α'. In turn, inasmuch as the observer wishes to reconcile their belief state with what they observe, they have an incentive to attempt to infer (consciously or not) why the speaker did indeed prefer u to α'. Furthermore, it stands to reason that the more favorable α'

appears to be relative to *u* prima facie in the context of utterance, the more likely it is to receive such consideration, ceteris paribus—for the more favorable α′ appears to be relative to *u*, the more explanation is needed concerning why it was not used instead.

We may distill these very general dynamics to the very general principles in 22. (Note: for present purposes, I intend to use the term ALTERNATIVE utterance in the broadest possible sense: something that might have been said instead of the observed utterance.)

> (22) Let *u* be an utterance issued by speaker *S* and heard by hearer *H* in context *C*, and suppose that *H* assumes that *S* attempted to optimize the cost-benefit ratio of *u*. Then, insofar as *H* wishes to reconcile their own beliefs with the speaker's utterance:
>   a. RECONCILIATION: If, given an alternative utterance α, it is not obvious to *H* why *S* would have preferred *u* to α in *C*, *H* will have reason to try to infer why *S* used *u* and not α.[12]
>   b. WEIGHING ALTERNATIVES: The better the mix of costs and benefits an alternative α appears to *H* to have relative to *u* in *C*, the more likely *H* is to draw an inference concerning *S*'s nonuse of α, ceteris paribus.

The intent of 22 here is to lay out in the most general terms why language users enrich utterances vis-à-vis alternatives and why some alternatives are more likely than others to figure in such enrichment. Such principles, as we will see, are the very foundation of the core social meanings of interest in the present work.

At the same time, however, 22 does not itself tell the whole story, largely because it does not speak to the question of what aspects of utterances tend to be viewed as beneficial or costly. Fortunately, previous pragmatic research offers a great deal in answering that very question. For Grice (1975), for instance, trying to optimize an utterance's costs and benefits more specifically means attempting to select the maximally 'cooperative' utterance—that is, one that is based on sound evidence (QUALITY), is appropriately informative (QUANTITY and RELATION), and is clear and concise (MANNER). Considering a classic case, suppose a speaker says 'some' in a context in which 'all' would have been relevant. Given that 'all' is just as relevant and more informative in the context without being any less clear or concise, it may appear prima facie to be more cooperative (to offer a better cost-benefit ratio) than 'some'. By 22, this naturally raises the question of why the speaker did not say 'all'. The speaker is assumed to be trying to provide the 'best' (in this case, maximally cooperative) utterance, so there must be some reason for their avoiding 'all'. One leading possibility is that saying 'all' would have meant saying something the speaker believed to be false, violating the quality maxim—hence the possible 'not all' inference.

Similarly, in subsequent pragmatic research the benefit of an utterance is most typically taken to be the amount of relevant information it encodes, and, on the assumption that we try to conserve effort (in terms of retrieval, production, and processing), costs are usually determined in terms of the utterance's markedness (Horn 1984, Levinson 2000, Sperber & Wilson 2004, inter alia; see Horn 2004 for a useful review). As I show in §5, this perspective provides the key ingredients needed to explain the core meanings of interest in this work.

In the meantime, it is worth noting that this focus on what Grice (1975:47) calls 'a maximally effective exchange of information' has generally also meant focusing on nonentailed meanings that directly enrich an utterance's entailments concerning the sit-

---

[12] I should also add here the condition that the hearer believes that *u* and α are distinct enough from each other that using one over the other would have nontrivial consequences.

uation being described and the identity of the referents involved (though see research on politeness, e.g. Lakoff 1973, Brown & Levinson 1987 et seq., for exceptions).

But of course conveying propositional, descriptive content is not the only function of talking. Concerning the effects of interest here, for instance, we find a different case, where the potential inferences are more focused on characterizing discourse participants (in this case, the speaker) than on enriching the description of the situation under discussion (cf. Potts's 2007 conception of EXPRESSIVE content). Consider again example 1a, repeated here with one of its potential implicata as 23.

(23) **The Americans** love cars. +> The speaker is not an American or wishes to express distance from Americans.

Here, the implicatum does not narrow the interpretation of the term *Americans* to some specific kind of American or lead to a more specific interpretation of *love* or *cars*. Rather, it says something about the speaker—more specifically, about the speaker's relation to the group under discussion. Other recent studies (e.g. Davis & Potts 2010, Acton & Potts 2014, Glass 2015, Beltrama & Staum Casasanto 2017) similarly focus on meanings that speak to discourse participants' traits, views, or relations.

Happily, the general pragmatic perspective discussed in this section, together with the specific principles worked out in previous pragmatic research, can in fact take us very far in explaining both kinds of nonentailed meaning, as I show in §5. At the same time, doing justice to the nonentailed meanings of *the*-plurals—not to mention the range of nonentailed meanings at large—means, for one thing, more thoroughly incorporating our understanding of variation, indexical meanings (see below), context-dependence, discourse norms, and social and linguistic ideologies. These domains have been matters of intense focus in sociolinguistics.

4.2. A VIEW FROM SOCIOLINGUISTICS. Like pragmatics, sociolinguistics has long recognized nonentailed meaning. Within sociolinguistics, however, the nonentailed meanings of interest, like the meanings of interest here, generally concern discourse participants' traits, stances, and relations. The primary focus has been on INDEXICAL meaning (Peirce 1955, Silverstein 2003, Eckert 2008)—essentially, a given form's associations and connotations—and its relation to language variation and change. This sense of *indexical* is different from but related to the use of the term in reference to semantically context-sensitive expressions like *I*, *here*, and *now*. In essence, forms accumulate associations with features of the contexts in which they are observed, including information concerning speakers' stances, traits, and the like. The form then 'points' to these stances, traits, and so on, hence the term *indexical*.

Now when a form bears indexical meaning, it can then be used to send signals tied to that meaning. Different individuals, occupying different places in social space and having different goals, are variably likely to desire to send particular signals, and therefore to use particular forms to varying degrees. The indexical meaning of a form thus helps explain variation in its use (Eckert 2008, 2012). And with each new use a form accrues new associations, thereby shifting its significance and social utility.

Variationist research rooted in this perspective—what Eckert (2012) calls the 'third wave' of variationist research—in fact has much in common at a foundational level with the pragmatic perspectives discussed above. Third-wave studies, like pragmatic studies, treat language users as purposive agents who speak in service of their goals, consistent with 21 and 22 (see e.g. Burnett 2017). As Eckert (2012:97–98) puts it, 'patterns of variation do not simply unfold from the speaker's structural position in a system of production, but are part of the active—stylistic—production of social differentiation', as

observed, for example, in Kiesling's (1998) study of 'fraternity brothers using apical variants of *-ing* to invoke power' (Eckert 2012:95–96).

Both pragmatics and third-wave variationism thus seek to document and explain how language users employ semiotic resources to serve their goals. That said, third-wave variationism offers distinct perspectives and insights to the present study and to pragmatic research more generally. First and perhaps foremost, as a matter of defining principle, third-wave variationism foregrounds the connection between differences in meaning and patterns of use, as in the study presented in this work. Second, while pragmatics research often (though certainly not without exception) focuses on relatively robust nonentailed meanings that tend 'to go through unless a special context is present' (Horn 2004:5), for variationist sociolinguistics, context and variation are the sine qua non. Accordingly, the latter has greatly illuminated the thoroughgoing context-sensitivity of nonentailed meaning. Campbell-Kibler (2007), for instance, examining the English *-ing* suffix, found systematic variation in individuals' interpretations of the *-in'* variant, which can suggest relaxedness, laziness, or even pretentiousness, depending on features of the speaker's utterance, the addressee's beliefs, the speaker's traits, and so on. Similarly, Podesva (2007:489) shows that while the social meaning of falsetto may center on a 'core expressive meaning', its particularized meaning and function on a given occasion depend heavily on the linguistic and extralinguistic context. As demonstrated above and discussed further in §6, this is certainly the case with *the*-plurals, which may figure in an act of othering in one case (as in 6), lend an air of objectivity to the speaker in another (as in the MGC), and have more to do with presenting the group of interest as a bloc than with separating the speaker from that group in yet another (as in 2, 18, and 19).

Third-wave variationism also complements pragmatics in its focus on indexical meanings. As I suggest below concerning loaded expressions like *the blacks*, even where a given nonentailed meaning is largely rooted in context-insensitive aspects of meaning, indexicality still has a role to play in explaining why particular instances of particular expressions get particular, distinctive interpretations.

With these perspectives in mind, in §5 I present an account of the phenomenon of interest here. The basic strategy will be to recruit insights from pragmatics to derive the core social meanings of interest—namely the distancing and monolithizing effect of using *the*-plurals—and explain why these meanings are less likely to arise with BPs. Then, in §6, drawing on insights from the variationist sociolinguistics tradition, I discuss the crucial role of context and indexicality in how *the*-plurals are interpreted.

Before turning to the account, it bears emphasis that I do not assume that the account presents procedures that language users consciously follow in real time. Rather, at a minimum, the account presents a rationality-based explanation of why the social meanings of interest tend to obtain. I leave as an open question the degree to which the meanings are already indexicalized.

5. Deriving the core meanings of interest.

5.1. Which alternatives? In accordance with the pragmatic perspective outlined in §4.1, the core meanings of interest will be derived by considering *the*-plurals vis-à-vis alternative expressions that one might otherwise have thought might offer the speaker a better mix of costs and benefits. The first question in deriving the core meanings of interest, then, is which potential alternatives appear to offer rich benefits at low costs relative to *the*-plurals.

As noted in §4.1, previous pragmatic research has gotten a great deal of mileage out of analyzing benefits in terms of the amount of relevant information an expression en-

codes and costs in terms of markedness. With that in mind, the very general principle in 22b typically takes on the form of some version of the more specific (neo-)Gricean principle in 24.

(24) WEIGHING ALTERNATIVES, V2 ((NEO-)GRICEAN): Given an utterance *u* issued by a speaker *S* in a context *C*: A hearer is likely to draw an inference concerning the speaker's nonuse of α insofar as it appears that α might offer a better mix of relevant information (benefit) and markedness (cost).

In the spirit of 24, it stands to reason that much research in the (neo-)Gricean tradition has centered on alternatives that, given an observed utterance *u*, encode strictly more relevant information than *u* but are no (or not considerably) more marked, and those that are strictly less marked but encode no (or not considerably) less relevant information. This is the nature of Levinson's 2000 Q-PRINCIPLE and M-PRINCIPLE and Katzir's 2007 CONVERSATIONAL PRINCIPLE, for instance, and it serves our present purposes well, as I show shortly.

What I have referred to here in general terms as markedness is operationalized in different ways depending on the theory. For Katzir (2007), it is operationalized as structural complexity, so that given any two expressions φ and ψ, ψ is no more complex than φ if and only if 'we can transform φ into ψ by a finite series of deletions, contractions, and replacements of constituents in φ with constituents of the same category' (Katzir 2007:679), and ψ is strictly less complex than φ if the quoted criterion is satisfied but the reverse is not.[13] Considering a DP of the form *the Xs*, this means that any grammatical expression of the form (Det) + *Xs* will be no more marked than *the Xs*. Of particular interest here, this includes expressions of the form *we Xs* (*we*-plurals) and the BP *Xs*, the latter of which is strictly less marked. As I show below, the availability of these alternatives is central to the social meanings of interest.

Other theories offer more complex notions of markedness that include additional considerations. Levinson's (2000) theory, for instance, includes not only structural complexity but also (in)frequency, among other things. There is indeed a good argument to be made that *we*-plurals, being relatively infrequent, are at least somewhat more marked than *the Xs*. In any case, because of the role of structural complexity, in Levinson's system there is still a wider gap in markedness between BPs and *we*-plurals than between *the*-plurals and *we*-plurals. In turn, thinking in terms of the markedness side of 24, on both Levinson's theory and Katzir's theory we should expect that *we*-plurals would sooner figure in the interpretation of *the*-plurals than in the interpretation of BPs.

As for informativity, on the interpretations of interest neither *we*-plurals nor BPs are any less informative than *the*-plurals. As noted at the outset, the uses of *the Xs* of interest are those for which the expression is used to say something about all or typical instantiations of a particular group or kind of individuals. In such statements, BPs and *the*-plurals can generally be used without leading to any dramatic differences in truth conditions, as in 25.

(25) a. **Americans** do love cars!
     b. **The Americans** do love cars!

Both examples in 25 attribute car loving to all or typical Americans, both allow for some exceptions, and so on. Furthermore, assuming an all-or-typical interpretation, the paral-

---

[13] Though see Katzir 2007:n. 11 on making allowances for additional markedness considerations.

lel example in 26 is no less specific in its truth conditions, and as I discuss further in §5.2, it is in fact more informative than the examples in 25 in terms of presuppositions.[14]

(26) **We Americans** do love cars!

Taking all of this together, what we have in *Xs* and *we Xs* are alternatives to *the Xs* that are, respectively: (i) strictly less marked without being less informative, and (ii) strictly more informative without being hugely (or, on Katzir's 2007 account, any) more marked.

There are, of course, other alternatives one might consider that would fit the same bill. Nevertheless, given the nature of the effect, I do not discuss here such alternatives (e.g. generic demonstratives; Bowdle & Ward 1995), though the availability of these expressions as alternatives does add interesting subtleties to the picture. For more on the relations between these various kinds of expression, see Acton 2014.

In the spirit of the Gricean principles in §4.1, the core meanings of interest are derived by considering why a speaker might prefer to use a *the*-plural over a BP or a *we*-plural. Let us therefore take a closer look at what sets these expressions apart from each other semantically, with an eye toward identifying the potential advantages and disadvantages that come with using them.

5.2. THE ENCODED MEANING AND FORM OF THE RELEVANT EXPRESSIONS.

THE. The work of *the* in its plural uses is to pick out particular collections of object-level individuals.[15] Following Sharvy (1980), Wolter (2006), Elbourne (2008, 2013), and Acton (2014), inter alia, I take the semantics of *the* to be a function that takes a property *P* and returns the function that maps a given situation *s* to the maximum indi-

---

[14] It is worth pointing out that, in addition to having all-or-typical readings, 25b and 26 are amenable to readings for which the subjects denote a particular proper subset of Americans. One referee notes, for instance, that '(ia) and (ib) could be uttered by an individual at an international conference to communicate that just the Americans at the conference are arrogant; (ic), on the other hand, can only be heard as a statement about Americans in general'.

(i) a. **The Americans** are arrogant.
   b. **We Americans** are arrogant.
   c. **Americans** are arrogant.

Thus, *the*-plurals and *we*-plurals, where an all-or-typical interpretation is intended, are susceptible to misinterpretation in a way that BPs are not. If, following Grice (1975) and his manner maxim, one takes an utterance's utility to vary directly with its perspicuity (ceteris paribus, and assuming a desire to communicate clearly), one could thus further argue that where the goal is to make an all-or-typical statement, the susceptibility of *the*-plurals and *we*-plurals to this kind of misinterpretation adds to their markedness/undesirability relative to BPs. If this is so, it only further strengthens the account of the distancing and monolithizing effects of *the*-plurals developed in §5.4, which relies on *the*-plurals being more marked than BPs in all-or-typical statements. In any case, the account does not depend on there being this extra source of markedness for *the*-plurals, since, as already established, they are already more marked than BPs given their greater structural complexity. See §5.3 for additional discussion.

[15] Although its object-level use is overwhelmingly more common, *the* can also be used in referring to TAXONOMIC KINDS. Such reference is (as the name suggests) highly associated with scientific discourse and resists predication involving nonessential properties (Dayal 2004). Talk about humans as taxonomic kinds thus often borders on comical or condescending, as in (i).

(i) The American (a most curious creature) loves cars.

Such uses, like the relevant uses of *the*-plurals, also suggest distance between the speaker and the relevant group—but, given the relative markedness of the former and their associations with scientific discourse, they are arguably even more distancing.

Also on the topic of *the*-singulars, a referee points out that object-level uses of *the* with proper names like *the Google* or *the Facebook* can suggest a ' "naïve older person" persona'. Whether this is best attributed to misnaming relatively recent technology or to something about *the* specifically (or both) must await future work.

vidual satisfying *P* in *s* (where defined). For plurals, this means the plural individual comprised of all atomic individuals satisfying *P* in *s* (Link 1983)—or, equivalently for our purposes, the set of all individuals satisfying *P* in *s*—as a collective. For instance, the encoded meaning of *the dogs*, where defined, maps a given situation *s* to the collection of all dogs in *s*.

Importantly, in picking out particular, well-defined collections of individuals as a unit, *the*-plurals draw boundary lines in the domain of discourse and, in turn, can foreground questions of who is and is not part of the group in question.

First-person forms. First-person forms also definitively pick out object-level individuals. Unlike *the*-plurals, however, the speech-act agent (generally, the speaker) figures directly in their propositional meaning. As a first approximation, *we* maps a situation to the speaker and some contextually relevant group of associated individuals in that situation. Complicating things only slightly is the fact that *we*, like *the*, can take an NP sister, as in *we Americans*. Moreover, Nunberg (1993) points out some important intricacies associated with *we* that suggest a need for a more complex formulation. Generalizing to the 'worst case', one can say that, where defined, *we* takes an optional property *P* argument and returns a function that maps a situation *s* to the collection of all individuals in *s* that satisfy *P* and some contextually determined property borne by the speech-act agent (nearly always the speaker) (Acton 2014).

Thus, assuming an all-or-typical interpretation, both *the*-plurals and *we*-plurals ultimately contribute the same collection of individuals to the proposition being expressed, but *we*-plurals are strictly more informative than *the*-plurals in their presuppositions; only the former presuppose that the speech-act agent instantiates the relevant property. As explained more fully below, in this contrast we have the germ of the distancing effect of *the*-plurals. The central question, however, is not only why *the*-plurals tend to invite such an inference, but also why they are more likely to do so than BPs, and why they are more likely to present the relevant group of individuals as a monolith. This takes us to the encoded meaning of BPs.

Bare plurals. The crucial semantic difference between BPs and *the*-plurals for our purposes is that unlike *the*-plurals, BPs do not definitively pick out particular collections of object-level individuals. In turn, they neither draw definitive dividing lines in the domain of discourse, nor foreground who is in or out of a particular group.

Let us consider the three basic types of BP interpretations (e.g. Farkas & de Swart 2007), as exemplified in 27–29.

(27) Dogs aren't extinct.   Kind-level

(28) Dogs bark when threatened.   Characterizing

(29) Dogs are barking.   Episodic   (Dayal 2013, ex. 5c)

Kind-level interpretations, like *extinct* in 27, involve predicates that 'do not apply to regular individuals like John and Fido, but only to species' (Farkas & de Swart 2007:1659). That is, such BPs pick out kinds of things, rather than the things (object-level individuals) themselves. Indeed, on Chierchia's (1998) and Dayal's (2004) accounts, object-level individuals never enter the picture on kind-level interpretations; rather, the interpretation comes from the application of a kind operator to the property encoded by the BP.

Characterizing sentences, like 28, differ from sentences like 27 in that their main predicate is not kind-level, but object-level. Many semantic accounts of characterizing sentences involve abstract quantification over individuals or situations via a generic operator, so that the denotation of 28 is something like 30, where Gen is the relevant operator (see e.g. Papafragou 1996 for discussion).

(30) $\text{G\scriptsize EN}_{s,x}\ (dog'(x)(s)\ \&\ threatened'(x)(s))\ (barks'(x)(s))$ (cf. Farkas & de Swart 2007)

The interpretation of 28, then, is something like 'Typically, in a situation involving a threatened dog, the dog will bark'. While various accounts of characterizing sentences differ in their details, they are all alike in that the BP ultimately contributes a property to the proposition. Thus, BPs do not deliver a well-defined collection of object-level individuals in characterizing sentences either.

Lastly are episodic sentences, like 29, which, rather than making a general claim about typical instantiations of the relevant kind or property, describe a particular situation. On most accounts of sentences like 29 (e.g. Chierchia 1998, Krifka 2003), the sentence ends up with an existential interpretation—so that 29 is interpreted as something like 'There are some dogs barking'. Even on Dobrovie-Sorin's (2009) analysis, where bare plurals denote sums of individuals, the precise makeup of the relevant sum remains indeterminate, so that *dogs* in 29 'refers' not to a particular, fully determined sum of individual dogs, but rather to SOME nonminimal, nonmaximal sum of individual dogs.

Worth specific mention are those episodic BP sentences that lack an indefinite feel, such as 31, for which 33 seems a better paraphrase than 32.

(31) Californians voted three-to-one in favor of the measure.
(32) There were some Californians who voted three-to-one in favor of the measure.
(33) Among Californians who voted, three in four voted in favor of the measure.

Krifka (2003:198–99) discusses similar examples, analyzing the relevant BPs as ultimately denoting kinds. Dayal (2013), by contrast, claims that an episodic sentence of the form BP + predicate is true in s just in case the predicate holds of some representative subpart of the collection of all individuals who instantiate the relevant kind in some situation s', where $s < s'$.

Again, on neither account do such BPs definitively pick out particular object-level individuals. Instead, they involve a lack of particularity, as evidenced in exchanges such as 34, where the respondent exploits this property in an act of evasion.

(34) Parent: And who is going to this party?
       Recalcitrant teenager: People!

On all of their various interpretations, then, BPs do not definitively pick out a collection of object-level individuals and are crucially less determinate than *the*-plurals.

5.3. THE MARKEDNESS OF DEFINITES IN ALL-OR-TYPICAL STATEMENTS. In each of the variants in 25 and 26, repeated here as 35 and 36, the truth conditions are in essence the same on an all-or-typical interpretation—all three convey that a given American will typically love cars, or something like that.

(35) a. **Americans** do love cars!
       b. **The Americans** do love cars!
(36) **We Americans** do love cars!

With regard to considerations of structural complexity (Levinson 2000, Katzir 2007, inter alia), it is no wonder that *the*-plurals (and *we*-plurals) are relatively marked in all-or-typical statements, given that BPs achieve the same basic truth-conditional ends through less costly means.[16] Indeed, in many cases *the*-plurals are so marked for all-or-

---

[16] A referee rightly points out that one might argue that, structurally speaking, BPs are as complex as *the*-plurals. The argument is that BPs include a phonologically null determiner in the form of a KIND type-shifter, GEN, or some other operator. However, even if this is the right structural analysis for BPs, the standard in the pragmatics literature, which I adopt here, is to treat overt morphological material as contributing more to the formal markedness of an expression than any potential covert material (see e.g. Horn 2004:16–17), and

typical statements that, unlike BPs, they strongly default to a particularized interpretation. Consider the minimal pairs in 37 and 38.

(37) a. Cats love tuna fish.
　　 b. The cats love tuna fish.
(38) a. I love cats!
　　 b. I love the cats!

In both pairs, the (a) sentences lend themselves to an all-or-typical interpretation, such that the speaker is saying something about cats in general. In contrast, the (b) sentences most naturally lend themselves to interpretations whereby the speaker is saying something about a particular collection of cats—the cats they own, the cats they are pet-sitting, or some other salient collection of cats. The same goes for a wide range of *the*-plurals; they tend to be evaluated relative to situations smaller than entire worlds or other macrosituations. (Consider: *books* vs. *the books*, *windows* vs. *the windows*, *electrons* vs. *the electrons*, etc.) That we default to a particularized reading for many *the*-plurals is perhaps unsurprising from a pragmatic perspective: it is reasonable for a hearer to suspect that a speaker using a *the*-plural intends to communicate something that a BP, which is less costly, could not or would not.

Of course, certain *the*-plurals are relatively amenable to all-or-typical interpretations—particularly those *the*-plurals whose constituent NPs correspond to groups whose members are frequently talked about in collective terms, such as sports teams (*the Detroit Tigers*), nations (*the Italians*), or political parties (*the Whigs*). Still, given the availability of less costly BPs, it is clear that *the*-plurals are relatively marked in all-or-typical statements.[17]

5.4. Deriving the core meanings of interest. Sections 2 and 3 provided extensive support for the claim that using a *the*-plural to talk about all or typical members of a group of individuals tends to depict that group as a monolith separate from the speaker, and to an extent that using a BP does not. The remaining question is why this should be.

To illustrate the dynamics involved, consider a scenario in which a group of attendees at an international conference are dining together. Bauer, a European, remarks on the great number of automobiles in the United States. Jones, whose nationality is unknown (Jones's accent suggests Canadian or US citizenship), responds with either 35a or 35b. Let us take each in turn, beginning with 35b.

Scenario one: *the americans*. In this case, the speaker has used 35b (*The Americans do love cars!*). Being a relatively marked choice, it may raise the question of why it was preferred to other potential utterances. As discussed in §5.1, (neo-)Gricean research furnishes 35a (*Americans*) and 36 (*we Americans*) as two clear candidates for alternatives that might appear to offer a better mix of costs and benefits, and with good reason: 35a because it is no less informative than the actual utterance while being considerably less marked, and 36 because it is more informative in its presuppositions (entailing that the speech-act agent is an American) without being much more marked. Returning to the

the same goes for optimality-theoretic research on the nominal domain such as de Swart & Zwarts 2009 (I thank this same referee for pointing me to this latter work). Hence, even if they are indeed comparable in structural complexity, BPs are still less formally marked than *the*-plurals.

[17] There is in addition an argument to be made that the fact that *the*-plurals, unlike BPs, can receive a particularized definite interpretation or an all-or-typical interpretation further adds to their markedness where an all-or-typical interpretation is intended. See n. 14. In any event, as noted there, the account developed here does not depend on there being this extra source of markedness for *the*-plurals.

general Gricean principle in 22a, the remaining question is why the speaker might nonetheless prefer 35b to those alternatives.

Concerning alternatives that encode additional relevant information but are not considerably (or, even better, any) more marked, the typical pragmatic line is that not using such alternatives signals that using them would come at the cost of saying something false or insufficiently evidenced (see e.g. Horn's 1984 and Levinson's 2000 Q-implicatures and Katzir's 2007 'conversational principle'). The spirit of the present account of the distancing effect is the same, but the focus here is on felicity rather than truth conditions. Let us assume that the hearers believe that Jones has full knowledge of her own nationality and that information about her nationality would be deemed relevant. With 35b, Jones has used a relatively marked form (cf. 35a), but not 36, which is more informative without being any more structurally marked. Following the line above, then, one obvious reason to forgo 36 would be because it would come at the cost of saying something 'false' or, more precisely, infelicitous. That is, a hearer might well conclude that Jones forwent the apparent benefits of saying 'we Americans' because it would have resulted in a presupposition failure—in particular, because the speaker is not an American.[18]

This particular calculation includes the assumption that the speaker would indeed find it beneficial to provide the information that she is an American, were it true. We could just as well relax the assumption of Gricean cooperativity and, in turn, the assumption that, modulo considerations of markedness, Jones will share all relevant information with her audience. In that case, a hearer might conclude not that Jones is not an American, but that, if she is, she does not think it would benefit her to share that information. Either way, given the availability of the unused alternative *we Americans*, we very naturally come to the potential inference that the speaker either is not an American or does not want to make it known that she is—hence the distancing effect.

This takes us to the other alternative of interest, 35a (*Americans*), which has essentially the same truth conditions as 35b on an all-or-typical interpretation. Encountering 'the Americans', one might reasonably ask why the speaker would incur the extra markedness cost of the definite determiner *the* if doing so does not materially change the truth conditions of the utterance. What benefit might a *the*-plural offer that would make the optional determiner worth using?

One potential answer goes back to the consideration just discussed: the availability of a less marked, determiner-free BP where a *the*-plural is used highlights the NONuse of another determiner—namely *we*—the latter of which specifically encodes the speech-

---

[18] Christopher Kennedy (p.c.) notes that *the*-plurals seem to tend toward an interpretation that excludes the speaker not just in all-or-typical statements, but in general. He gives the example of a member of a university's Department of Linguistics and Philosophy saying (i), regarding a faculty meeting. Kennedy's (defeasible) default interpretation is that the speaker of (i) is a philosopher, rather than a linguist.

   (i) (I think) **the linguists** would prefer to hold the meeting in room 101.

As with the case of 35b, I claim that it is the competition between *the linguists* and first-person plurals like *we (linguists)* that yields this effect: *we (linguists)* is more informative in terms of its presuppositions (semantically including the speaker) without being more formally complex, thus potentially offering a better mix of costs and benefits on the face of things and, in turn, potentially raising the question of why the speaker opted for *the linguists* instead. Again, one plausible explanation is that using a first-person form would have involved the extra cost of saying something infelicitous—hence the potential inference that the speaker is not a linguist. Of course, the interpretation (i) would receive in context would depend on multiple factors, including the knowledge of the hearer concerning who belongs to which group. Also relevant here is the fact that, unlike in the case of all-or-typical statements, replacing the *the*-plural with a BP yields a very different interpretation, making BPs a less suitable alternative.

act agent's membership in the relevant group, along the lines just discussed. In this way, the availability of the BP *Americans* can strengthen the distancing effect.

But there is something else that 35b offers the speaker that 35a does not. Specifically, with its *the*-plural, it delivers the set of all Americans as a collective and foregrounds the boundary around that collective. Inasmuch as the speaker would like to present the individuals being discussed as a solid, well-defined group, this property of *the*-plurals offers a benefit to the speaker. Thus, encountering 35b, a hearer might reasonably entertain the notion that Jones deliberately incurred the extra markedness cost of a *the*-plural to avail herself of this benefit and depict Americans as a monolith—hence the monolithizing effect.

Again, unlike in prototypical cases of implicature, this effect is not so much about a speaker's desire to directly enrich the utterance's entailments concerning the situation being described or to help resolve reference. Rather, this effect, like the distancing effect, involves an inference concerning the speaker's view of and relation to the individuals under discussion. Nonetheless, the same general pragmatic principles underlying familiar cases like scalar implicature explain these social meanings.

SCENARIO TWO: *AMERICANS*. In this case, the result is rather different. Let us first consider the distancing effect. In the previous scenario, this effect was born of the contrast between the observed form and *we Americans*. Here, however, the observed form is relatively unmarked, so the gap in markedness between *we Americans* and the observed form is considerably greater. Thus, whereas vis-à-vis *the Americans* there is good reason to ask why the speaker did not opt for the more informative and structurally similar *we Americans*, here the larger gap in markedness provides a solid explanation in and of itself: the speaker may have avoided *we Americans* simply because it was not worth the considerable extra markedness cost. In turn, a hearer is less likely to conclude that saying 'we Americans' would have been infelicitous or that the speaker is deliberately withholding information, and, in turn, the distancing effect is less likely.

As for the monolithizing effect, again the gap in markedness between *the Americans* and *Americans* itself provides a rationale for opting for the latter over the former—leaving nothing to be explained. Moreover, even if in some context *the Americans* were considered a relevant alternative to an observed use of *Americans*, the contrast would not engender a monolithizing effect. For in that case, the observed form is the one that does NOT delimit a collective of individuals. If anything, opting for *Americans* would signal that the speaker does not view the individuals being talked about as a monolith.

Thus we have a principled explanation for the distancing and monolithizing effects of *the*-plurals and for why these effects are less likely to show up with BPs. The dynamics are rooted in general pragmatic principles and contrasts in the markedness and meaning of BPs, *the*-plurals, and first-person forms.

The effect, then, is hardly accidental. Indeed, thinking crosslinguistically, this account predicts that we should observe the same sort of effect in languages where BPs are the default form for all-or-typical statements but for which, at least in certain circumstances, a definite plural may be so used. My Swedish- and Dutch-speaking consultants confirm that this is the case for their native languages, so that the (b) sentences in 39 and 40 are more likely than the (a) sentences to signal distance from the relevant group.

(39) Swedish
    a. Norsk-ar   gillar   skidåkning.
       Norwegian-PL like.PRS skiing
         'Norwegians like skiing.'

   b. Norsk-ar-na   gillar   skidåkning.
     Norwegian-PL-DEF like.PRS skiing
       'The Norwegians like skiing.'

(40) Dutch
   a. Belg-en   zijn   dol   op friet.
     Belgian.M-PL be.PRS.PL crazy on fries
       'Belgians love fries.'
   b. De   Belg-en   zijn   dol   op friet.
     the.NOM.PL Belgian.M-PL be.PRS.PL crazy on fries
       'The Belgians love fries.'

By the same token, in languages where a definite marker is required in all-or-typical statements, we should not expect to find so strong a distancing effect from phrases headed by that marker, because in those languages there is not a briefer, less marked alternative available.

6. Context and related considerations.

6.1. Multiple possible interpretations. Having said all of that, while the pragmatic approach taken here derives the core social meanings of interest, it does not tell the full story. Perhaps most obviously, as research in third-wave variationist sociolinguistics has made abundantly clear, the precise use to which a linguistic form is put and the precise interpretation it actually receives varies from one context to the next. As noted already, a *the*-plural may indeed signal speaker nonmembership, as in 4 (repeated here as 41), but in other cases it may have little at all to do with distancing and more to do with presenting the group of interest as a unified front, as in 2, 18, and 19 (repeated here as 42–44). That is, the availability of more than one viable alternative (in this case, BPs and first-person forms) means that there is more than one possible explanation for a speaker's use of a *the*-plural—either or both of which may be in play in a particular instance.

(41) Paid for by Bernie 2016 (not **the billionaires**).
                  (from US Senator Bernie Sanders's presidential campaign website)
(42) **The Tea Party Patriots** depend on grassroots activism where it counts: near you. Local Groups are the backbone of **our** activities across the nation.
                       (from the website of the 'Tea Party Patriots')
(43) **The Democrats** are united on the need for a new direction in Iraq.
                       (John Larson (D-CT), 26 June 2006)
(44) What **the Republicans** are going to stand against is tying the funds our soldiers need to do their jobs to benchmarks thought up by special interest groups.   (Adrian Smith (R-NE), 9 March 2007)

Thus, as in Podesva's (2007) study of the social meaning of falsetto, the ultimate interpretation depends importantly on what the purpose of the speech event is taken to be, what the discourse participants already know about the speaker's group membership and attitudes, and so on.

   As another example, in the case of the MGC, where presenting an air of journalistic objectivity is particularly important, first-person plural forms are more marked and therefore more costly than in other situations. Therefore, ceteris paribus, we should expect that using a *the*-plural in the MGC would be less likely to raise the question of why the speaker did not use a first-person form than it would in another context where first-person forms are not so costly. In turn, we should also predict that the inference that the speaker is not a member of the relevant group would be less likely in the MGC, ceteris paribus.

   Thus, even if we concern ourselves only with semantic differences between the relevant expressions, there will still be context-based variation in how a given *the*-plural

will be interpreted. Then there is the issue of the role of indexicality and ideology, in the sense discussed in §4.2, to which I now turn.

6.2. Distancing and derogation. Using a form that signals one's nonmembership or downplays one's membership in a group is a way of putting distance between oneself and that group. It is tempting to assume that signaling distance in turn means derogation. In general, it seems that we tend to mark distance between ourselves and things with which we do not wish to be associated.

But sociolinguistic research on the role of indexicality, context, and ideology in interpretation (such as the work cited in §4.2) cautions against assuming that such links will hold across all contexts. After all, derogation does not simply logically follow from distancing. In fact, marking distance is also a way of showing respect or deference. Silverstein's (2003) treatment of *tu/vous* distinctions, for instance, centers on the notion of distance, as do other accounts of formality in language. Another particularly instructive case concerns the expression *the wife*, used to refer to one's own spouse (or, in a parallel case, the wife of an addressee). To my ear, this form—which, being in competition with forms like *my wife*, clearly signals distance between speaker and spouse—has a distinctly derogatory tone. Though not entailed, this association between distance and derogation certainly figures in my interpretation of *the wife*. In contrast, however, I have spoken with others who claim that, in their social circles, use of this phrase is just as likely to convey deference to one's spouse as it is to convey pejoration.

Now if the social meaning of *the wife* were based strictly on reasoning over encoded meaning, then, holding grammar fixed, we would expect the social meaning to be roughly the same regardless of the speech community and situation in which it was being used. But this is not the case. It must be, then, that the differences in interpretation have some other source. Drawing on the sociolinguistics literature, the clearest candidates are differences in ideology and differences in previous experiences with the form (Eckert 2008), which lead to differences in indexical meaning from one person to the next.

The same holds for the derogation associated with the use of certain *the*-plurals. Marking distance does not entail a negative stance. Instead, it seems that signaling a negative stance with a *the*-plural, though certainly related to distancing, comes in part from ideology or associations with prior uses. Presumably, such is the case with loaded expressions like *the gays* and *the blacks*, which have repeatedly figured in statements of pejoration and marginalization.

Thus the analysis of *the Americans* and *Americans* presented in §5 is only part of the picture—a rather large and important part, but a proper part nonetheless. The precise interpretation of any particular BP or *the*-plural in a particular instance depends importantly on the expression's indexical meaning for the speech participants, which, as a crucial part of what the form brings to the table, enters into the dynamics outlined in §5 just as well as semantic content does. If an addressee has repeatedly heard a form in pejorative statements, one can bet that those pejorative associations will color their interpretation of the use at hand. So, too, will their ideologies and many other factors besides.

Still, considering the encoded descriptive meaning of *the*-plurals and relevant alternatives in the light of general pragmatic principles circumscribes the range of expected interpretations in a principled way. Given the dynamics discussed above, we should be very surprised to find in any community that *the*-plurals are commonly used to signal special intimacy between the speaker and the group under discussion, for instance.

7. Conclusion. The perspective taken herein thus offers principled explanations of nonentailed meanings, in the tradition of pragmatics, while at the same time underscoring the role of context and linking patterns of variation to meaning, in the tradition of third-wave variationist research. This kind of perspective can be applied to a wide range

of phenomena (Acton 2014). Indeed, a similar theme can be found running through a broadening array of recent studies on meanings concerning interlocutors' traits, moods, stances, and relations—covering other determiners (Davis & Potts 2010, Acton & Potts 2014), modals (Glass 2015), and intensifiers (Beltrama & Staum Casasanto 2017).

Eckert (2008) rightly points out that context, variation, and indexicality are inextricably bound. This work presents a picture whereby encoded semantic meaning is part of that same web of interrelations. The pragmatic principles used to explain reference resolution and enrichments to situation descriptions can be extended to illumine the nature of social meaning; the very differences in meaning that one may be tempted to gloss over in a variationist study can in fact help explain the nature of the variation of interest.

Practically speaking, on the present view theories of semantic meaning, together with pragmatic theory, provide a basis for predictions about social meanings and patterns of variation. By the same token, observed social meanings and patterns of variation provide data for semantic theories to reckon with. Eckert (2011:4–5) is right: 'The bottom line is that it is time to integrate the study of variation with the study of meaning in language more generally'. Semantics, pragmatics, and sociolinguistics need each other.

REFERENCES

ABBOTT, BARBARA. 2008. Issues in the semantics and pragmatics of definite descriptions in English. *Reference: Interdisciplinary perspectives*, ed. by Jeanette Gundel and Nancy Hedberg, 61–72. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780195 331639.003.0003.

ACTON, ERIC K. 2014. *Pragmatics and the social meaning of determiners*. Stanford, CA: Stanford University dissertation. Online: http://purl.stanford.edu/sk702hj6068.

ACTON, ERIC K., and CHRISTOPHER POTTS. 2014. That straight talk: Sarah Palin and the sociolinguistics of demonstratives. *Journal of Sociolinguistics* 18.3–31. DOI: 10.1111 /josl.12062.

BATES, DOUGLAS; MARTIN MÄCHLER; BEN BOLKER; and STEVE WALKER. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67.1–48. DOI: 10 .18637/jss.v067.i01.

BELTRAMA, ANDREA, and LAURA STAUM CASASANTO. 2017. *Totally* tall sounds *totally* younger: Intensification at the socio-semantics interface. *Journal of Sociolinguistics* 21.154–82. DOI: 10.1111/josl.12230.

BIRNER, BETTY, and GREGORY WARD. 1994. Uniqueness, familiarity, and the definite article in English. *Berkeley Linguistics Society* 20.93–102. DOI: 10.3765/bls.v20i1.1479.

BOWDLE, BRIAN F., and GREGORY WARD. 1995. Generic demonstratives. *Berkeley Linguistics Society* 21.32–43. DOI: 10.3765/bls.v21i1.1396.

BROWN, PENELOPE, and STEPHEN C. LEVINSON. 1987. *Politeness: Some universals in language use*. Cambridge: Cambridge University Press.

BURNETT, HEATHER. 2017. Sociolinguistic interaction and identity construction: The view from game-theoretic pragmatics. *Journal of Sociolinguistics* 21.238–71. DOI: 10.1111 /josl.12229.

CAMERON, RICHARD, and SCOTT SCHWENTER. 2013. Pragmatics and variationist sociolinguistics. *The Oxford handbook of sociolinguistics*, ed. by Robert Bayley, Richard Cameron, and Ceil Lucas, 464–83. Oxford: Oxford University Press. DOI: 10.1093 /oxfordhb/9780199744084.013.0023.

CAMPBELL-KIBLER, KATHRYN. 2007. Accent, (ing), and the social logic of listener perceptions. *American Speech* 82.32–64. DOI: 10.1215/00031283-2007-002.

CHESHIRE, JENNY. 2005. Syntactic variation and beyond: Gender and social class variation in the use of discourse-new markers. *Journal of Sociolinguistics* 9.479–508. DOI: 10 .1111/j.1360-6441.2005.00303.x.

CHIERCHIA, GENNARO. 1998. Reference to kinds across languages. *Natural Language Semantics* 6.339–405. DOI: 10.1023/A:1008324218506.

CLIFT, ELEANOR. 2008. The Hillary I know. *Newsweek*, March 8, 2008. Online: http://www .newsweek.com/cliftthe-hillary-i-know-84083.

COPPOCK, ELIZABETH, and DAVID BEAVER. 2012. Weak uniqueness: The only difference between definites and indefinites. *Proceedings of Semantics and Linguistic Theory (SALT)* 22.527–44. DOI: 10.3765/salt.v22i0.2651.

DAVIES, MARK. 2008–. *The Corpus of Contemporary American English: 520 million words, 1990–present*. Online: http://corpus.byu.edu/coca/.

DAVIS, CHRISTOPHER, and CHRISTOPHER POTTS. 2010. Affective demonstratives and the division of pragmatic labor. *Logic, language, and meaning: 17th Amsterdam Colloquium revised selected papers*, ed. by Maria Aloni, Harald Bastiaanse, Tikitu de Jager, and Katrin Schulz, 42–52. Berlin: Springer. DOI: 10.1007/978-3-642-14287-1_5.

DAYAL, VENEETA. 2004. Number marking and (in)definiteness in kind terms. *Linguistics and Philosophy* 27.393–450. DOI: 10.1023/B:LING.0000024420.80324.67.

DAYAL, VENEETA. 2013. On the existential force of bare plurals across languages. *From grammar to meaning: The spontaneous logicality of language*, ed. by Ivano Caponigro and Carlo Cecchetto, 49–80. Cambridge: Cambridge University Press. DOI: 10.1017 /CBO9781139519328.006.

DE SWART, HENRIËTTE, and JOOST ZWARTS. 2009. Less form—more meaning: Why bare singular nouns are special. *Lingua* 119.280–95. DOI: 10.1016/j.lingua.2007.10.015.

DJALALI, ALEX J. 2013. *House Proceedings Corpus*. Stanford, CA: Stanford University.

DOBROVIE-SORIN, CARMEN. 2009. Existential bare plurals: From properties back to entities. *Lingua* 119.296–313. DOI: 10.1016/j.lingua.2007.10.016.

ECKERT, PENELOPE. 2008. Variation and the indexical field. *Journal of Sociolinguistics* 12. 453–76. DOI: 10.1111/j.1467-9841.2008.00374.x.

ECKERT, PENELOPE. 2011. The future of variation studies. Plenary talk given at New Ways of Analyzing Variation (NWAV) 40, Georgetown University. Online: https://web.stanford .edu/~eckert/PDF/FutureOFVariation.pdf.

ECKERT, PENELOPE. 2012. Three waves of variation study: The emergence of meaning in the study of variation. *Annual Review of Anthropology* 41.87–100. DOI: 10.1146/annurev -anthro-092611-145828.

ELBOURNE, PAUL. 2008. Demonstratives as individual concepts. *Linguistics and Philosophy* 31.409–66. DOI: 10.1007/s10988-008-9043-0.

ELBOURNE, PAUL. 2013. *Definite descriptions*. (Oxford studies in semantics and pragmatics.) Oxford: Oxford University Press.

FARKAS, DONKA F., and HENRIËTTE DE SWART. 2007. Article choice in plural generics. *Lingua* 117.1657–76. DOI: 10.1016/j.lingua.2006.06.011.

GLASS, LELIA. 2015. Strong necessity modals: Four socio-pragmatic corpus studies. *University of Pennsylvania Working Papers in Linguistics (Selected papers from New Ways of Analyzing Variation 43)* 21.79–88. Online: http://repository.upenn.edu/pwpl/vol21/iss2 /10/.

GOODMAN, NOAH D., and MICHAEL C. FRANK. 2016. Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Sciences* 20.818–29. DOI: 10.1016/j.tics .2016.08.005.

GRICE, H. PAUL. 1975. Logic and conversation. *Syntax and semantics, vol. 3: Speech acts*, ed. by Peter Cole and Jerry Morgan, 43–58. New York: Academic Press.

HORN, LAURENCE R. 1984. Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. *Meaning, form, and use in context: Linguistic applications*, ed. by Deborah Schiffrin, 11–42. Washington, DC: Georgetown University Press.

HORN, LAURENCE R. 2004. Implicature. In Horn & Ward, 3–28.

HORN, LAURENCE R., and GREGORY WARD (eds.) 2004. *Handbook of pragmatics*. Oxford: Blackwell.

HORNADAY, ANN. 2014. 'Pride' revisits labor and gay politics in '80s Britain. *Washington Post*, October 8, 2014. Accessed via *Infotrac Newsstand*, July 11, 2015.

JOHNSON, DANIEL EZRA. 2009. Getting off the GoldVarb standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and Linguistics Compass* 3.359–83. DOI: 10.1111/j.1749-818X.2008.00108.x.

KATZIR, RONI. 2007. Structurally-defined alternatives. *Linguistics and Philosophy* 30.669–90. DOI: 10.1007/s10988-008-9029-y.

KELLER, RUDI. 1994. *On language change: The invisible hand in language*, trans. by Brigitte Nerlich. London: Routledge.

Kiesling, Scott F. 1998. Men's identities and sociolinguistic variation: The case of fraternity men. *Journal of Sociolinguistics* 2.69–100. DOI: 10.1111/1467-9481.00031.

Krifka, Manfred. 2003. Bare NPs: Kind-referring, indefinites, both, or neither? *Proceedings of Semantics and Linguistic Theory (SALT)* 13.180–203. DOI: 10.3765/salt.v13i0.2880.

Kudlow, Larry. 2014. About Kudlow & Company LLC. *Kudlow & Company*. Online: http://kudlow.com/about.php, last accessed July 1, 2016.

Lakoff, Robin. 1973. The logic of politeness; or, minding your P's and Q's. *Chicago Linguistic Society* 9.292–305.

Levinson, Stephen C. 2000. *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, MA: MIT Press.

Levy, Roger. 2015. Grammatical knowledge is fundamentally probabilistic. Paper presented at the annual meeting of the Linguistic Society of America, Portland.

Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. *Meaning, use, and interpretation of language*, ed. by Rainer Baüerle, Christoph Schwarze, and Arnim von Stechow, 302–23. Berlin: Walter de Gruyter.

Maltin, Leonard. 2005. John McLaughlin. *Turner Classic Movies*. Online: http://www.tcm.com/tcmdb/person/128201|0/John-Mclaughlin/biography.html, last accessed July 1, 2016.

McAlley, John. 2009. Frank's 'Americans,' still revelatory after 50 years. *NPR.org*, February 13, 2009. Online: http://www.npr.org/2009/02/13/100673221/franks-americans-still-revelatory-after-50-years, last accessed April 14, 2014.

McLaughlin, John (Producer). 1998–2001. *The McLaughlin Group*. Chicago: WTTW National Productions.

Nunberg, Geoffrey. 1993. Indexicality and deixis. *Linguistics and Philosophy* 16.1–43. DOI: 10.1007/BF00984721.

O'Donnell, Lawrence. 2011. Rewrite. *The Last Word*, MSNBC, August 1, 2011. Online: http://www.nbcnews.com/id/21134540/#43980204, last accessed July 1, 2016. (Also cited on *Wikipedia*.)

Outsiders' chance. 2016. *The Economist*, January 30, 2016. Online: https://www.economist.com/briefing/2016/01/30/outsiders-chance, last accessed July 2, 2016.

Page, Clarence. 2014. Page's page. *The Chicago Tribune*. Online: http://www.chicagotribune.com/news/opinion/page/, last accessed May 20, 2014.

Papafragou, Anna. 1996. On generics. *UCL Working Papers in Linguistics* 8.1–35.

Peirce, Charles Sanders. 1955. Logic as semiotic: The theory of signs. *Philosophical writings of Peirce*, ed. by Justus Buchler, 98–119. New York: Dover Books.

Placher, William C. 2009. How does Jesus save? *Christian Century*. In Davies 2008–.

Podesva, Robert J. 2007. Phonation type as a stylistic variable: The use of falsetto in constructing a persona. *Journal of Sociolinguistics* 11.478–504. DOI: 10.1111/j.1467-9841.2007.00334.x.

Potts, Christopher. 2007. The expressive dimension. *Theoretical Linguistics* 33.165–97. DOI: 10.1515/TL.2007.011.

R Core Team. 2012. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. Online: http://www.R-project.org/.

Republicans backed into a corner; Alleged spy spills Chinese secrets to U.S.; A win for Sandusky? 2012. *Erin Burnett OutFront*. In Davies 2008–.

Roberts, Craige. 2002. Demonstratives as definites. *Information sharing: Reference and presupposition in language generation and interpretation*, ed. by Kees van Deemter and Roger Kibble, 89–136. Stanford, CA: CSLI Publications.

Robertson, Campbell. 2014. Among whites, protests stir a range of emotions and a lot of perplexity. *New York Times*, August 22, 2014. Accessed via *Infotrac Newsstand*, July 11, 2015.

Romaine, Suzanne. 1984. On the problem of syntactic variation and pragmatic meaning in sociolinguistic theory. *Folia Linguistica* 18.409–38. DOI: 10.1515/flin.1984.18.3-4.409.

Saad, Lydia. 2015. U.S liberals at record 24%, but still trail conservatives. *Gallup*, January 9, 2015. Online: http://www.gallup.com/poll/180452/liberals-record-trail-conservatives.aspx, last accessed July 2, 2016.

Sankoff, David. 1988. Sociolinguistics and syntactic variation. *Linguistics: The Cambridge survey. Vol. 4: Language: The socio-cultural context*, ed. by Frederick J. Newmeyer, 140–61. Cambridge: Cambridge University Press.

Sharvy, Richard. 1980. A more general theory of definite descriptions. *Philosophical Review* 89.607–24. DOI: 10.2307/2184738.

Silverstein, Michael. 2003. Indexical order and the dialectics of sociolinguistic life. *Language & Communication* 23.193–229. DOI: 10.1016/S0271-5309(03)00013-2.

Sperber, Dan, and Deirdre Wilson. 2004. Relevance theory. In Horn & Ward, 607–32.

Strawson, P. F. 1950. On referring. *Mind* 59.320–44. Online: https://www.jstor.org/stable/2251176.

Tagliamonte, Sali A., and Derek Denis. 2014. Expanding the transmission/diffusion dichotomy: Evidence from Canada. *Language* 90.90–136. DOI: 10.1353/lan.2014.0016.

Traugott, Elizabeth Closs. 2011. Pragmatics and language change. *The Cambridge handbook of pragmatics*, ed. by Keith Allan and Kasia M. Jaszczolt, 549–66. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139022453.030.

Walker, James A. 2010. *Variation in linguistic systems*. New York: Routledge.

Wolter, Lynsey Kay. 2006. *That's that: The semantics and pragmatics of demonstrative noun phrases*. Santa Cruz: University of California, Santa Cruz dissertation.

Woodward, Richard B. 2009. 'The Americans' revisited. *The Wall Street Journal*, November 18, 2009. Online: https://www.wsj.com/articles/SB1000142405274870432200457447761913139506, last accessed April 14, 2014.

Department of English Language and Literature
Eastern Michigan University
612 Pray-Harrold
Ypsilanti, MI 48197
[eacton1@emich.edu]