Chameleons in imagined conversations: A new approach to understanding coordination of linguistic style in dialogs

Cristian Danescu-Niculescu-Mizil and Lillian Lee
Department of Computer Science, Cornell University
cristian@cs.cornell.edu, llee@cs.cornell.edu

Abstract

Conversational participants tend to immediately and unconsciously adapt to each other's language styles: a speaker will even adjust the number of articles and other function words in their next utterance in response to the number in their partner's immediately preceding utterance. This striking level of coordination is thought to have arisen as a way to achieve social goals, such as gaining approval or emphasizing difference in status. But has the adaptation mechanism become so deeply embedded in the language-generation process as to become a reflex? We argue that fictional dialogs offer a way to study this question, since authors create the conversations but don't receive the social benefits (rather, the imagined characters do). Indeed, we find significant coordination across many families of function words in our large movie-script corpus. We also report suggestive preliminary findings on the effects of gender and other features; e.g., surprisingly, for articles, on average, characters adapt more to females than to males.

1 Introduction

"...it is dangerous to base any sociolinguistic argumentation on the evidence of language in fictional texts only" (Bleichenbacher (2008), crediting Mareš (2000))

The chameleon effect is the "nonconscious mimicry of the postures, mannerisms, facial expressions, and other behaviors of one's interaction partners" (Chartrand and Bargh, 1999).[1] For example, if one conversational participant crosses their arms, their partner often unconsciously crosses their arms as well. The effect occurs for language, too, ranging from matching of acoustic features such as accent, speech rate, and pitch (Giles et al., 1991; Chartrand and van Baaren, 2009) to lexico-syntactic priming across adjacent or nearby utterances (Bock, 1986; Pickering and Garrod, 2004; Ward and Litman, 2007; Reitter et al., 2011).

Our work focuses on adjacent-utterance coordination with respect to classes of function words. To exemplify the phenomenon, we discuss two short conversations.

• *First example:* The following exchange from the movie "The Getaway" (1972) demonstrates quantifier coordination.

Doc: At least you were outside.
Carol: It doesn't make much difference where you are [...]

Note that "Carol" used a quantifier, one that is different than the one "Doc" employed. Also, notice that "Carol" could just as well have replied in a way that doesn't include a quantifier, for example, "It doesn't really matter where you are...".

• *Second example:* Levelt and Kelter (1982) report an experiment involving preposition coordination. Shopkeepers who were called and asked "At what time does your shop close?" were significantly more likely to say "At five o'clock" than "five o'clock".[2]

[1] The term is a reference to the movie *Zelig*, wherein a "human chameleon" uncontrollably takes on the characteristics of those around him. The term is meant to contrast with "aping", a word connoting intentional imitation.

Related terms include adaptation, alignment, entrainment, priming, and Du Bois' dialogic syntax.

[2] This is an example of lexical matching manifested as part of syntactic coordination.

Coordination of function-word class has been previously documented in several settings (Niederhoffer and Pennebaker, 2002; Taylor and Thomas, 2008; Ireland et al., 2011; Gonzales et al., 2010), the largest-scale study being on Twitter (Danescu-Niculescu-Mizil et al., 2011).

Problem setting People don't consciously track function words (Levelt and Kelter, 1982; Segalowitz and Lane, 2004; Petten and Kutas, 1991) — it's not easy to answer the question, "how many prepositions were there in the sentence I just said?". Therefore, it is quite striking that humans nonetheless instantly adapt to each other's function-word rates. Indeed, there is active debate regarding what mechanisms cause nonconscious coordination (Ireland et al., 2011; Branigan et al., 2010).

One line of thought is that convergence represents a social strategy[3] whose aim is to gain the other's social approval (Giles, 2008; Street and Giles, 1982) or enhance the other's comprehension (Clark, 1996; Bortfeld and Brennan, 1997).[4] This hypothesis is supported by studies showing that coordination is affected by a number of social factors, including relative social status (Natale, 1975; Gregory and Webster, 1996; Thakerar et al., 1982) and gender role (Bilous and Krauss, 1988; Namy et al., 2002; Ireland and Pennebaker, 2010).

But an important question is whether the adaptation mechanism has become so deeply embedded in the language-generation process as to have transformed into a reflex not requiring any social triggering.[5] Indeed, it has been argued that unconscious mimicry is partly innate (Chartrand and Bargh, 1999), perhaps due to evolutionary pressure to foster relationships (Lakin et al., 2003).

To answer this question, we take a radical approach: we consider a setting in which the persons *generating* the coordinating dialog are different from those *engaged* in the dialog (and standing to reap the social benefits) — imagined conversations, specifically, scripted movie dialogs.

Life is beautiful, but cinema is paradise A priori, it is not clear that movie conversations would exhibit convergence. Dialogs between movie characters are not truthful representations of real-life conversations. They often are "too carefully polished, too rhythmically balanced, too self-consciously artful" (Kozloff, 2000), due to practical and artistic constraints and scriptwriting practice (McKee, 1999). For example, mundane phenomena such as stuttering and word repetitions are generally nonexistent on the big screen. Moreover, writers have many goals to accomplish, including the need to advance the plot, reveal character, make jokes as funny as possible, and so on, all incurring a cognitive load.

So, the question arises: do scripted movie dialogs, in spite of this quasi-artificiality and the aforementioned generation/engagement gap, exhibit the real-life phenomenon of stylistic convergence? When imagining dialogs, do scriptwriters (nonconsciously[6]) adjust the respondent's replies to echo the initiator's use of articles, prepositions, and other apparently minor aspects of lexical choice? According to our results, this is indeed the case, which has fascinating implications.

First, this provides evidence that coordination, assumed to be driven by social motivations, has become so deeply embedded into our ideas of what conversations "sound like" that the phenomenon occurs even when the person generating the dialog is not the recipient of the social benefits.[7]

Second, movies can be seen as a controlled environment in which preconceptions about the relation between communication patterns and the social features of the participants can be studied. This gives us the opportunity to understand how people (scriptwriters) nonconsciously *expect* convergence to relate to factors such as gender, status and relation type. Are female characters thought to accommodate more to male characters than vice-versa?

Furthermore, movie scripts constitute a corpus that is especially convenient because meta-features

[3] In fact, social signaling may also be the evolutionary cause of chameleons' color-changing ability (Stuart-Fox et al., 2008).

[4] For the purpose of our discussion, we are conflating social-approval and audience-design hypotheses under the category of *social strategy*.

[5] This hypothesis relates to characterizations of alignment as an unmediated mechanism (Pickering and Garrod, 2004).

[6] The phenomenon of real-life language convergence is not widely known among screenplay authors (Beth F. Milles, professor of acting and directing, personal communication).

[7] Although some writers may perhaps imagine themselves "in the shoes" of the recipients, recall that authors generally don't include in their scripts the repetitions and ungrammaticalities of "real-life" speech.

like gender can be more or less readily obtained.

Contributions We check for convergence in a corpus of roughly 250,000 conversational exchanges from movie scripts (available at `http://www.cs.cornell.edu/~cristian/movies`). Specifically, we examine the set of nine families of stylistic features previously utilized by Ireland et al. (2011), and find a statistically significant convergence effect for all these families. We thereby provide evidence that language coordination is so implanted within our conception of conversational behavior that, even if such coordination is socially motivated, it is exhibited even when the person generating the language in question is not receiving any of the presumed social advantages.

We also study the effects of gender, narrative importance, and hostility. Intriguingly, we find that these factors indeed "affect" movie characters' linguistic behavior; since the characters aren't real, and control of stylistic lexical choice is largely nonconscious, the effects of these factors can only be springing from patterns existing in the scriptwriters' minds.

Our findings, by enhancing our understanding of linguistic adaptation effects in stylistic word choice and its relation to various socially relevant factors, may in the future aid in practical applications. Such an understanding would give us insight into how and what kinds of language coordination yield more satisfying interactions — convergence has been already shown to enhance communication in organizational contexts (Bourhis, 1991), psychotherapy (Ferrara, 1991), care of the mentally disabled (Hamilton, 1991), and police-community interactions (Giles et al., 2007). Moreover, a deeper understanding can aid human-computer interaction by informing the construction of natural-language generation systems, since people are often more satisfied with encounters exhibiting appropriate linguistic convergence (Bradac et al., 1988; van Baaren et al., 2003), even when the other conversational participant is known to be a computer (Nass and Lee, 2000; Branigan et al., 2010).

2 Related work not already mentioned

Linguistic style and human characteristics Using stylistic (i.e., non-topical) elements like articles and prepositions to characterize the utterer in some way has a long history, including in authorship attribution (Mosteller and Wallace, 1984; Juola, 2008), personality-type classification (Argamon et al., 2005; Oberlander and Gill, 2006; Mairesse et al., 2007), gender categorization (Koppel et al., 2002; Mukherjee and Liu, 2010; Herring and Paolillo, 2006), identification of interactional style (Jurafsky et al., 2009; Ranganath et al., 2009), and recognizing deceptive language (Hancock et al., 2008; Mihalcea and Strapparava, 2009).

Imagined conversations There has been work in the NLP community applying computational techniques to fiction, scripts, and other types of text containing imagined conversations. For example, one recent project identifies conversational networks in novels, with the goal of evaluating various literary theories (Elson et al., 2010; Elson and McKeown, 2010). Movie scripts were used as word-sense-disambiguation evaluation data as part of an effort to generate computer animation from the scripts (Ye and Baldwin, 2006). Sonderegger (2010) employed a corpus of English poetry to study the relationship between pronunciation and network structure. Rayson et al. (2001) computed part-of-speech frequencies for imaginative writing in the British National Corpus, finding a typology gradient progressing from conversation to imaginative writing (e.g., novels) to task-oriented speech to informative writing. The data analyzed by Oberlander and Gill (2006) consisted of emails that participants were instructed to write by imagining that they were going to update a good friend on their current goings-on.

3 Movie dialogs corpus

To address the questions raised in the introduction, we created a large set of imagined conversations, starting from movie scripts crawled from various sites.[8] Metadata for conversation analysis and duplicate-script detection involved mostly-automatic matching of movie scripts with the IMDB movie database; clean-up resulted in 617 unique titles tagged with genre, release year, cast lists, and

[8] The source of these scripts and more detail about the corpus are given in the README associated with the Cornell movie-dialogs corpus, available at `http://www.cs.cornell.edu/~cristian/movies`.

IMDB information. We then extracted 220,579 conversational exchanges between pairs of characters engaging in at least 5 exchanges, and automatically matched these characters to IMDB to retrieve gender (as indicated by the designations "actor" or "actress") and/or billing-position information when possible (≈9000 characters, ≈3000 gender-identified and ≈3000 billing-positioned). The latter feature serves as a proxy for narrative importance: the higher up in the credits, the more important the character tends to be in the film.

To the best of our knowledge, this is the largest dataset of (metadata-rich) imaginary conversations to date.

4 Measuring linguistic style

For consistency with prior work, we employed the nine LIWC-derived categories (Pennebaker et al., 2007) deemed by Ireland et al. (2011) to be processed by humans in a generally non-conscious fashion. The nine categories are: articles, auxiliary verbs, conjunctions, high-frequency adverbs, impersonal pronouns, negations, personal pronouns, prepositions, and quantifiers (451 lexemes total).

It is important to note that language coordination is multimodal: it does not necessarily occur simultaneously for all features (Ferrara, 1991), and speakers may converge on some features but diverge on others (Thakerar et al., 1982); for example, females have been found to converge on pause frequency with male conversational partners but diverge on laughter (Bilous and Krauss, 1988).

5 Measuring convergence

Niederhoffer and Pennebaker (2002) use the correlation coefficient to measure accommodation with respect to linguistic style features. While correlation at first seems reasonable, it has some problematic aspects in our setting (we discuss these problems later) that motivate us to employ an alternative measure.

We instead use a convergence measure introduced in Danescu-Niculescu-Mizil et al. (2011) that quantifies how much a given feature family t serves as an *immediate trigger* or stimulus, meaning that one person's utterance exhibiting such a feature triggers the appearance of that feature in the respondent's immediate reply.

For example, we might be studying whether one person A's inclusion of articles in an utterance triggers the usage of articles in respondent B's reply. Note that this differs from asking whether B uses articles more often when talking to A than when talking to other people (it is not so surprising that people speak differently to different audiences). This also differs from asking whether B eventually starts matching A's behavior in later utterances within the same conversation. We specifically want to know whether each utterance by A triggers an *immediate* change in B's behavior, as such instantaneous adaptation is what we consider the most striking aspect of convergence, although immediate and long-term coordination are clearly related.

We now describe the statistic we employ to measure the extent to which person B accommodates to A. Consider an arbitrary conversational exchange started by A, and let a denote A's initiating utterance and $b_{\hookrightarrow a}$ denote B's reply to a.[9] Note that we use lowercase to emphasize when we are talking about individual utterances rather than all the utterances of the particular person, and that thus, the arrow in $b_{\hookrightarrow a}$ indicates that we mean the reply to the specific single utterance a. Let a^t be the indicator variable for a exhibiting t, and similarly for $b^t_{\hookrightarrow a}$. Then, we define the convergence $Conv_{A,B}(t)$ of B to A as:

$$P(b^t_{\hookrightarrow a} = 1 | a^t = 1) - P(b^t_{\hookrightarrow a} = 1). \quad (1)$$

Note that this quantity can be negative (indicating *divergence*). The overall degree $Conv(t)$ to which t serves as a trigger is then defined as the expectation of $Conv_{A,B}(t)$ over all initiator-respondent pairs:

$$Conv(t) \stackrel{def}{=} E_{\text{pairs}(A,B)}(Conv_{A,B}(t)). \quad (2)$$

Comparison with correlation: the importance of asymmetry[10] Why do we employ $Conv_{A,B}$, Equation (1), instead of the well-known correlation coefficient? One reason is that correlation fails to

[9] We use "initiating" and "reply" loosely: in our terminology, the conversation ⟨A: "Hi." B: "Eaten?" A: "Nope."⟩ has two exchanges, one initiated by A's "Hi", the other by B's "Eaten?".

[10] Other asymmetric measures based on conditional probability of occurrence have been proposed for adaptation within monologues (Church, 2000) and between conversations (Stenchikova and Stent, 2007). Since our focus is different, we control for different factors.

capture an important asymmetry. The case where $a^t = 1$ but $b^t_{\hookrightarrow a} = 0$ represents a true failure to accommodate; but the case where $a^t = 0$ but $b^t_{\hookrightarrow a} = 1$ should not, at least not to the same degree. For example, a may be very short (e.g., "What?") and thus not contain an article, but we don't assume that this completely disallows B from using articles in their reply. In other words, we are interested in whether the presence of t acts as a trigger, not in whether $b_{\hookrightarrow a}$ exhibits t if and only if a does, the latter being what correlation detects.[11]

It bears mentioning that since a^t and $b^t_{\hookrightarrow a}$ are binary, a simple calculation shows that the covariance[12] $cov(a^t, b^t_{\hookrightarrow a}) = Conv_{A,B}(t) \cdot P(a^t = 1)$. But, the two terms on the right hand side are not independent: raising $P(a^t = 1)$ could cause $Conv_{A,B}(t)$ to decrease by affecting the first term in its definition, $P(b^t_{\hookrightarrow a} = 1 | a^t = 1)$ (see eq. 1).

6 Experimental results

6.1 Convergence exists in fictional dialogs

For each ordered pair of characters (A, B) and for each feature family t, we estimate equation (1) in a straightforward manner: the fraction of B's replies to t-manifesting A utterances that themselves exhibit t, minus the fraction of all replies of B to A that exhibit t.[13] Fig. 1 compares the average values of these two fractions (as a way of putting convergence values into context), showing positive differences for all of the considered families of features (statistically significant, paired t-test $p < 0.001$); this demonstrates that movie characters do indeed converge to each other's linguistic style on all considered trigger families.[14]

[11] One could also speculate that it is easier for B to (unconsciously) pick up on the presence of t than on its absence.

[12] The covariance of two random variables is their correlation times the product of their standard deviations.

[13] For each t, we discarded pairs of characters where some relevant count is < 10, e.g., where B had fewer than 10 replies manifesting the trigger.

[14] We obtained the same qualitative results when measuring convergence via the correlation coefficient, doing so for the sake of comparability with prior work (Niederhoffer and Pennebaker, 2002; Taylor and Thomas, 2008).

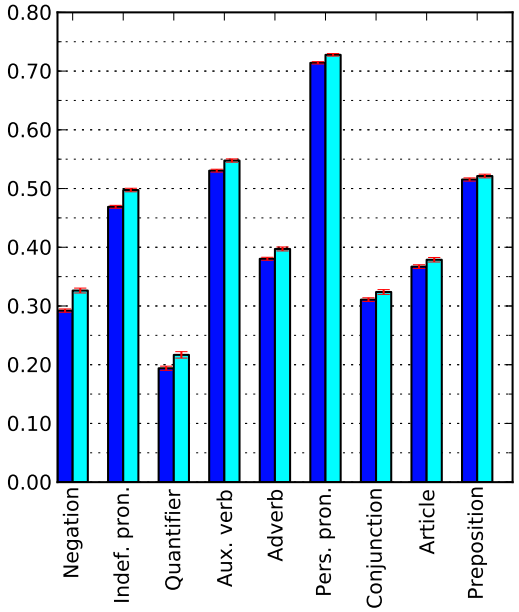


Figure 1: Implicit depiction of convergence for each trigger family t, illustrated as the difference between the means of $P(b^t_{\hookrightarrow a} = 1 | a^t = 1)$ (right/light-blue bars) and $P(b^t_{\hookrightarrow a} = 1)$ (left/dark-blue bars). (This implicit representation allows one to see the magnitude of the two components making up our definition of convergence.) The trigger families are ordered by decreasing convergence. All differences are statistically significant (paired t-test). In all figures in this paper, error bars represent standard error, estimated via bootstrap resampling (Koehn, 2004). (Here, the error bars, in red, are very tight.)

Movies vs. Twitter One can ask how our results on movie dialogs correspond to those for real-life conversations. To study this, we utilize the results of Danescu-Niculescu-Mizil et al. (2011) on a large-scale collection of Twitter exchanges as data on real conversational exchanges. Figure 2 depicts the comparison, revealing two interesting effects. First, Twitter users coordinate more than movie characters on all the trigger families we considered, which does show that the convergence effect is stronger in actual interchanges. On the other hand, from the perspective of potentially using imagined dialogs as proxies for real ones, it is intriguing to see that there is generally a correspondence between how much convergence occurs in real dialogs for a given feature family and how much convergence occurs for that feature in imagined dialogs, although conjunctions and articles show a bit less convergence in fictional

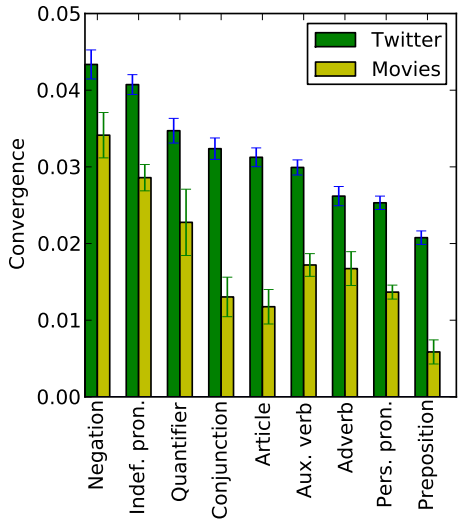


Figure 2: Convergence in Twitter conversations (left bars) vs. convergence in movie dialogs (right bars; corresponds to the difference between the two respective bars in Fig. 1) for each trigger family. The trigger families are ordered by decreasing convergence in Twitter.

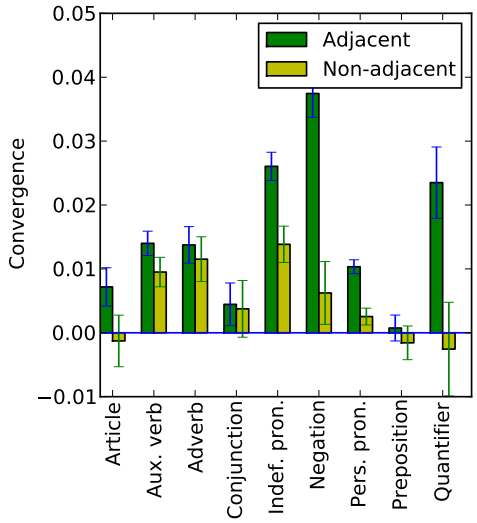


Figure 3: Immediate vs. within-conversation effects (for conversations with at least 5 utterances). Suppose that we have a conversation $a_1\, b_2\, a_3\, b_4\, a_5 \ldots$. The lefthand/dark-green bars show the usual convergence measure, which involves the utterance pair a_1 and b_2. The righthand/mustard-green bars show convergence based on pairs like a_1 and b_4 — utterances in the same conversation, but not adjacent. We see that there is a much stronger triggering effect for immediately adjacent utterances.

exchanges than this pattern would suggest.

6.2 Potential alternative explanations

Immediate vs. within-conversation effects An additional natural question is, how much are these accommodation effects due to an immediate triggering effect, as opposed to simply being a by-product of utterances occurring within the same conversation? For instance, could the results be due just to the topic of the conversation?

To answer this question requires measuring "convergence" between utterances that are not adjacent, but are still in the same conversation. To this end, we first restricted attention to those conversations in which there were at least five utterances, so that they would have the structure $a_1\, b_2\, a_3\, b_4\, a_5 \ldots$. We then measure convergence not between adjacent utterances, like a_1 and b_2, but where we skip an utterance, such as the pair a_1, b_4 or b_2, a_5. This helps control for topic effects, since b_4 and a_1 are still close and thus fairly likely to be on the same subject.[15]

Figure 3 shows that the level of convergence always falls off after the skipped utterance, sometimes dramatically so, thus demonstrating that the level of immediate adaptation effects we see cannot be solely explained by the topic of conversation or other conversation-level effects. These results accord with the findings of Levelt and Kelter (1982), where interposing "interfering" questions lowered the chance of a question's preposition being echoed by the respondent, and Reitter et al. (2006), where the effects of structural priming were shown to decay quickly with the distance between the priming trigger and the priming target.

Towards the same end, we also performed randomization experiments in which we shuffled the order of each participant's utterances in each conversation, while maintaining alternation between speakers. We again observed drop-offs in this randomized condition in comparison to immediate convergence, the main focus of this paper.

Self-coordination Could our results be explained entirely by the author converging to their own self, given that self-alignment has been documented

[15] It is true that they might be on different topics, but in fact even b_2 might be on a different subject from a_1.

(Pickering and Garrod, 2004; Reitter et al., 2006)? If that were the case, then the *characters* that the author is writing about should converge to themselves no more than they converge to different characters. But we ran experiments showing that this is not the case, thus invalidating this alternative hypothesis. In fact, characters converge to themselves much more than they converge to other characters.

6.3 Convergence and imagined relation

We now analyze how convergence patterns vary with the type of relationship between the (imagined) participants. Note that, given the multimodal character of convergence, treating each trigger family separately is the most appropriate way to proceed, since in past work, for the same experimental factor (e.g., gender), different features converge differently (refer back to §4). For clarity of exposition, we discuss in detail only the results for the *Articles* feature family; but the results for all trigger families are summarized in Fig. 7, discussed later.

Imagined gender Fig. 4(a) shows how convergence on article usage depends on the gender of the initiator and respondent. Females are more influential than males: movie characters of either gender accommodate more to female characters than to male characters (compare the **F**emale **init**iator bar with the **M**ale **init**iator bar, statistically significant, independent t-test, $p < 0.05$). Also, female characters seem to accommodate slightly more to other characters than male characters do (though not statistically significantly so in our data).

We also compare the amount of convergence between all the possible types of gendered initiator-respondent pairs involved (Fig. 4(b)). One can observe, for example, that male characters adapt less in same-gender situations (**M**ale-**M**ale conversations) than in mixed-gender situations (**F**emale initiator-**M**ale respondent), while the opposite is true for female characters (**F**emale-**F**emale vs. **M**ale-**F**emale).

Interpreting these results lies beyond the scope of this paper. We note that these results could be a correlate of many factors, such as the roles that male and female characters are typically assigned in movie scripts.[16]

Narrative importance Does the relative importance bestowed by the scriptwriter to the characters affect the amount of linguistic coordination he or she (nonconsciously) embeds in their dialogs? Fig. 5 shows that, on average, the lead character converges to the second-billed character more than vice-versa (compare left bar in **1st resp.** group with left bar in **2nd resp.** group).

One possible confounding factor is that there is significant gender imbalance in the data (82% of all lead characters are males, versus only 51% of the secondary characters). Could the observed difference be a direct consequence of the relation between gender and convergence discussed above? The answer is no: the same qualitative observation holds if we restrict our analysis to same-gender pairs (compare the righthand bars in each group in Fig. 5[17]).

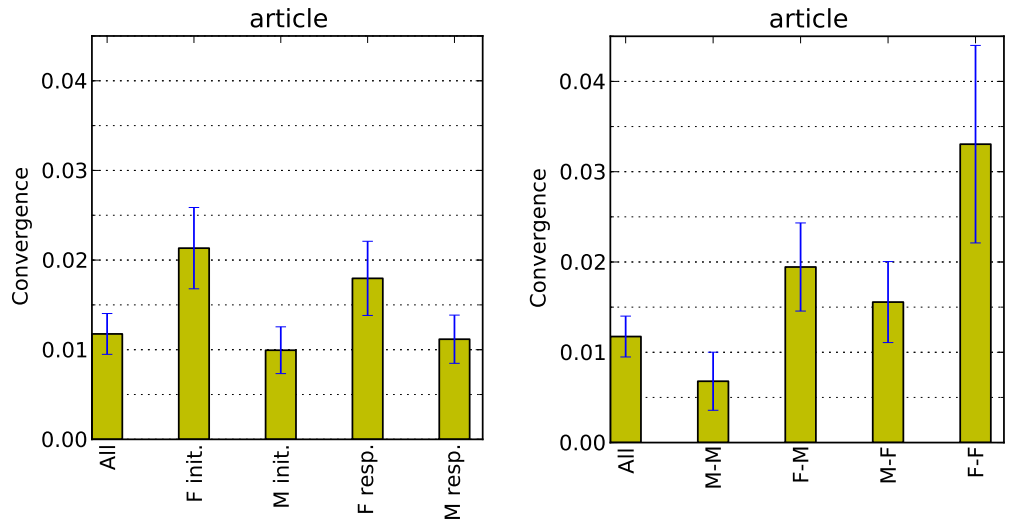
It would be interesting to see whether these results could be brought to bear on previous results regarding the relationship between social status and convergence, but such interpretation lies beyond the scope of this paper, since the connection between billing order and social status is not straightforward.

Quarreling The level of contention in conversations has also been shown to be related to the amount of convergence (Giles, 2008; Niederhoffer and Pennebaker, 2002; Taylor and Thomas, 2008). To test whether this tendency holds in the case of imagined conversations, as a small pilot study, we manually classified the conversations between 24 main pairs of characters from romantic comedies[18] as: *quarreling*, *some quarreling* and *no quarreling*. Although the experiment was too small in scale to provide statistical significance, the results (Fig. 6) suggest that indeed the level of convergence is affected by

[16] A comparison to previously reported results on real-life gender effects is not straightforward, since they pertain to different features; Ireland and Pennebaker (2010) show that females match their linguistic style more than males, where style matching is averaged over the same 9 trigger families we employ (they do not report gender effect for each family separately).

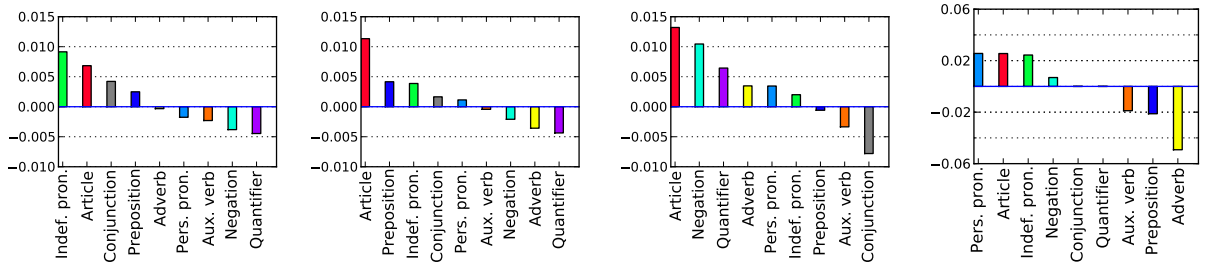
[17] Figure 5 also shows that our convergence measure does achieve negative values in practice, indicating divergence. Divergence is a rather common phenomenon which deserves attention in future work; see Danescu-Niculescu-Mizil et al. (2011) for an account.

[18] We chose the romantic comedy genre since it is often characterized by some level of contention between the two people in the main couple.



(a) (b)

Figure 4: Relation between *Article* convergence and imagined gender. (a) compares cases when the **init**iator and **resp**ondent are **M**ale or **F**emale; (b) compares types of gendered **initiator-respondent** relations: **M**ale-**M**ale, **F**emale-**M**ale, **M**ale-**F**emale, **F**emale-**F**emale. For comparison, the **All** bars represents the general *Article* convergence (illustrated in Fig. 1 as the difference between the two respective bars).



(a) **F resp.** minus **M resp.** (b) **F init.** minus **M init.** (c) **1st resp.** minus **2nd resp.** (d) **Quarrel** minus **No quarrel**

Figure 7: Summary of the relation between convergence and imagined gender (a and b), billing order (c), and quarreling (d). The bars represent the *difference* between the convergence observed in the respective cases; e.g., the **Article** (red) bar in (a) represents the difference between the **F resp.** and the **M resp.** bars in Fig. 4(a). In each plot, the trigger families are sorted according to the respective difference, but the color assigned to each family is consistent across plots. The scale of (d) differs from the others.

the presence of controversy: *quarreling* exhibited considerably more convergence for articles than the other categories (the same holds for personal and indefinite pronouns; see Fig. 7). Interestingly, the reverse is true for adverbs; there, we observe divergence for contentious conversations and convergence for non-contentious conversations (detailed plot omitted due to space constraints). This corresponds to Niederhoffer and Pennebaker's (2002) observations made on real conversations in their study of the Watergate transcripts: when the relationship between the two deteriorated, Richard Nixon converged more to John Dean on articles, but diverged on other features.[19]

Results for the other features Our results above suggest some intriguing interplay between convergence and gender, status, and level of hostility in imagined dialogs, which may shed light on how people (scriptwriters) nonconsciously *expect* con-

[19] Adverbs were not included in their study.

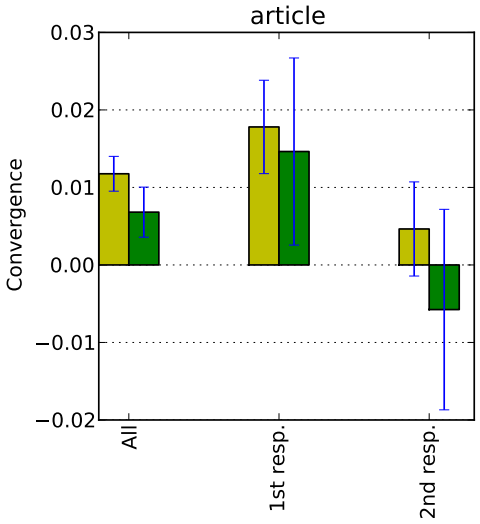


Figure 5: Comparison of the convergence of first-billed (lead) characters to second-billed characters (left bar in **1st resp.** group) to that of second-billed characters to leads (left bar in **2nd resp.** group); righthand bars (dark green) in each group show results for Male-Male pairs only.

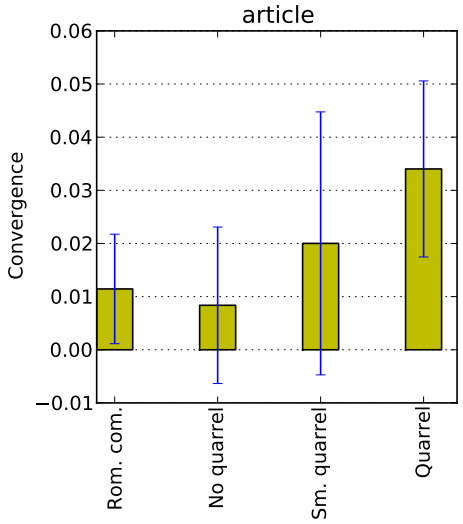


Figure 6: Relation between contention and convergence. The third bar combines *quarreling* and *some quarreling* to ameliorate data sparsity. For comparison, **Rom. com.** shows convergence calculated on all the conversations of the 24 romantic-comedy pairs considered in this experiment.

vergence to relate to such factors. (Interpreting these sometimes apparently counterintuitive findings is beyond the scope of this paper, but represents a fascinating direction for future work.) Fig. 7 shows how the nature of these relations depends on the trigger family considered. The variation among families is in line with the previous empirical results on the multimodality of convergence in real conversations, as discussed in §4.

7 Summary and future work

We provide some insight into the causal mechanism behind convergence, a topic that has generated substantial scrutiny and debate for over 40 years (Ireland et al., 2011; Branigan et al., 2010). Our work, along with Elson and McKeown (2010), advocates for the value of fictional sources in the study of linguistic and social phenomena. To stimulate such studies, we render our metadata-rich corpus of movie dialog public.

In §1, we described some practical applications of a better understanding of the chameleon effect in language; it boils down to improving communication both between humans and between humans and computers. Also, our results on contention could be used to further automatic controversy detection (Mishne and Glance, 2006; Gómez et al., 2008). Moreover, if we succeeded in linking our results on narrative importance to relative social status, we might further the development of systems that can infer social relationships in online social networks when conversational data is present but other, more explicit cues are absent (Wyatt et al., 2008; Bramsen et al., 2011). Such systems could be valuable to the rapidly expanding field of analyzing social networks.

Acknowledgments Many thanks for their help are due to Claire Cardie, Catalin Draghici, Susan Dumais, Shimon Edelman, Michael Gamon, Jon Kleinberg, Magdalena Narożniak, Alex Niculescu-Mizil, Myle Ott, Bo Pang, Morgan Sonderegger, plus NLP seminar participants Eric Baumer, Bin Lu, Chenhao Tan, Lu Wang, Bishan Yang, Ainur Yessenalina, and Marseille, and the anonymous reviewers (who went far beyond the call of duty!). Supported by NSF IIS-0910664, the Yahoo! FREP program, and a Yahoo! Key Scientific Challenges award.

References

Shlomo Argamon, Sushant Dhawle, and Moshe Koppel. 2005. Lexical predictors of personality type. *Proceedings of the 2005 Joint Annual Meeting of the Interface and the Classification Society of North America*.

Frances Bilous and Robert Krauss. 1988. Dominance and accommodation in the conversational behavior of same- and mixed-gender dyads. *Language and Communication*, 8:183–194.

Lukas Bleichenbacher. 2008. *Multilingualism in the movies: Hollywood characters and their language choices*. francke verlag, Jan.

J. Kathryn Bock. 1986. Syntactic persistence in language production. *Cognitive Psychology*, 18(3):355 – 387.

Heather Bortfeld and Susan E. Brennan. 1997. Use and acquisition of idiomatic expressions in referring by native and non-native speakers. *Discourse Processes*, 23(2):119–147.

Richard Y. Bourhis. 1991. Organizational communication and accommodation: Toward some conceptual and empirical links. In Howard Giles, Justine Coupland, and Nikolas Coupland, editors, *Contexts of Accommodation*. Cambridge University Press.

James J. Bradac, Anthony Mulac, and Ann House. 1988. Lexical diversity and magnitude of convergent versus divergent style shifting: Perceptual and evaluative consequences. *Language and Communication*, 8:213–228, Nov.

Philip Bramsen, Martha Escobar-Molana, Ami Patel, and Rafael Alonso. 2011. Extracting social power relationships from natural language. In *Proceedings of ACL HLT*.

Holly P. Branigan, Martin J. Pickering, Jamie Pearson, and Janet F. McLean. 2010. Linguistic alignment between people and computers. *Journal of Pragmatics*, 42(9):2355–2368.

Tanya L. Chartrand and John A. Bargh. 1999. The chameleon effect: The perception-behavior link and social interaction. *J. Personality and Social Psychology*, 76(6):893–910.

Tanya L. Chartrand and Rick van Baaren. 2009. Chapter 5: Human mimicry. In Mark P. Zanna, editor, *Advances in Experimental Social Psychology*, volume 41, pp. 219–274. Academic Press.

Kenneth W. Church. 2000. Empirical estimates of adaptation: the chance of two noriegas is closer to p/2 than p2. In *Proceedings of COLING*, pp. 180–186.

Herbert H. Clark. 1996. *Using language*. Cambridge University Press, second edition.

Cristian Danescu-Niculescu-Mizil, Michael Gamon, and Susan Dumais. 2011. Mark my words! Linguistic style accommodation in social media. In *Proceedings of WWW*.

David Elson and Kathleen McKeown. 2010. Automatic attribution of quoted speech in literary narrative. In *Proceedings of AAAI*.

David Elson, Nicholas Dames, and Kathleen McKeown. 2010. Extracting social networks from literary fiction. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, pp. 138–147.

Kathleen Ferrara. 1991. Accommodation in therapy. In *Accommodation theory: Communication, context, and consequences*. Cambridge University Press.

Howard Giles, Justine Coupland, and Nikolas Coupland. 1991. Accommodation theory: Communication, context, and consequences. In *Accommodation theory: Communication, context, and consequences*. Cambridge University Press.

Howard Giles, Michael Willemyns, Cynthia Gallois, and Michelle Anderson. 2007. Accommodating a new frontier: The context of law enforcement. In Klaus Fiedler, editor, *Social Communication*, Frontiers of Social Psychology, chapter 5, pp. 129–162.

Howard Giles. 2008. Communication accommodation theory. In *Engaging theories in interpersonal communication: Multiple perspectives*. Sage Publications.

Vicenç Gómez, Andreas Kaltenbrunner, and Vicente López. 2008. Statistical analysis of the social network and discussion threads in Slashdot. In *Proceedings of WWW*, pp. 645–654.

Amy L. Gonzales, Jeffrey T. Hancock, and James W. Pennebaker. 2010. Language style matching as a predictor of social dynamics in small groups. *Communication Research*, 37(1):3–19, Feb.

Stanford W. Gregory and Stephen Webster. 1996. A nonverbal signal in voices of interview partners effectively predicts communication accommodation and social status perceptions. *J. Personality and Social Psychology*, 70(6):1231–1240.

Heidi Hamilton. 1991. Accommodation and mental disability. In *Accommodation theory: Communication, context, and consequences*. Cambridge University Press.

Jeffrey T. Hancock, Lauren E. Curry, Saurabh Goorha, and Michael Woodworth. 2008. On lying and being lied to: A linguistic analysis of deception in computer-mediated communication. *Discourse Processes*, 45(1):1–23.

Susan C. Herring and John C. Paolillo. 2006. Gender and genre variation in weblogs. *Journal of Sociolinguistics*, Jan.

Walter Hill. 1972. *The Getaway*. Directed by Sam Peckinpah.

Molly E. Ireland and James W. Pennebaker. 2010. Language style matching in writing: Synchrony in essays, correspondence, and poetry. *J. Personality and Social Psychology*, 99(3):549–571.

Molly E. Ireland, Richard B. Slatcher, Paul W. Eastwick, Lauren E. Scissors, Eli J. Finkel, and James W. Pennebaker. 2011. Language style matching predicts relationship initiation and stability. *Psychological Science*, 22:39–44.

Patrick Juola. 2008. *Authorship Attribution*. Now Publishers.

Dan Jurafsky, Rajesh Ranganath, and Dan McFarland. 2009. Extracting social meaning: Identifying interactional style in spoken conversation. In *Proceedings of the NAACL*, pp. 638–646.

Philipp Koehn. 2004. Statistical significance tests for machine translation evaluation. *Proceedings of EMNLP*, pp. 388–395.

Moshe Koppel, Shlomo Argamon, and Anat Shimoni. 2002. Automatically categorizing written texts by author gender. *Literary and Linguistic Computing*.

Sarah Kozloff. 2000. *Overhearing Film Dialogue.* University of California Press.

Jessica L. Lakin, Valerie E. Jefferis, Clara Michelle Cheng, and Tanya L. Chartrand. 2003. The chameleon effect as social glue: Evidence for the evolutionary significance of nonconscious mimicry. *Journal of Nonverbal Behavior*, 27:145–162.

Willem J.M. Levelt and Stephanie Kelter. 1982. Surface form and memory in question answering. *Cognitive Psychology*, 14(1):78–106.

François Mairesse, Marilyn A. Walker, Matthias R. Mehl, and Roger K. Moore. 2007. Using linguistic cues for the automatic recognition of personality in conversation and text. *JAIR*, pp. 457–500.

Petr Mareš. 2000. Fikce, konvence a realita: K vícejazyčnosti v uměleckých textech [fiction, convention, and reality: On multilingualism in literary texts]. *Slovo a slovesnost*, 61(1):47–53.

Robert McKee. 1999. *Story: Substance, Structure, Style, and the Principles of Screenwriting.* Methuen.

Rada Mihalcea and Carlo Strapparava. 2009. The lie detector: Explorations in the automatic recognition of deceptive language. In *Proceedings of the ACL-IJCNLP 2009 Conference Short Papers*, pp. 309–312.

Gilad Mishne and Natalie Glance. 2006. Leave a reply: An analysis of weblog comments. *Third annual workshop on the Weblogging ecosystem.*

Frederick Mosteller and David L. Wallace. 1984. *Applied Bayesian and Classical Inference: The Case of the Federalist Papers*. Springer-Verlag.

Arjun Mukherjee and Bing Liu. 2010. Improving gender classification of blog authors. *EMNLP*.

Randall Munroe. 2010. http://xkcd.com/813/.

Laura L. Namy, Lynne C. Nygaard, and Denise Sauerteig. 2002. Gender differences in vocal accommodation. *J. Language and Social Psychology*, 21(4):422–432.

Clifford Nass and Kwan Min Lee. 2000. Does computer-generated speech manifest personality? An experimental test of similarity-attraction. In *Proceedings of CHI*, pp. 329–336.

Michael Natale. 1975. Convergence of mean vocal intensity in dyadic communication as a function of social desirability. *J. Personality and Social Psychology*, 32(5):790–804.

Kate G. Niederhoffer and James W. Pennebaker. 2002. Linguistic style matching in social interaction. *J. Language and Social Psychology*.

Jon Oberlander and Alastair J. Gill. 2006. Language with character: A stratified corpus comparison of individual differences in e-mail communication. *Discourse Processes*, 42(3):239–270.

James W. Pennebaker, Roger J. Booth, and Martha E. Francis. 2007. Linguistic inquiry and word count (LIWC): A computerized text analysis program. http://www.liwc.net/.

Cyma Van Petten and Marta Kutas. 1991. Influences of semantic and syntactic context on open- and closed-class words. *Memory and Cognition*, 19(1):95–112.

Martin Pickering and Simon Garrod. 2004. Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27(02):169–190.

Rajesh Ranganath, Dan Jurafsky, and Dan McFarland. 2009. It's not you, it's me: Detecting flirting and its misperception in speed-dates. In *Proceedings of EMNLP*, pp. 334–342.

Paul Rayson, Andrew Wilson, and Geoffrey Leech. 2001. Grammatical word class variation within the British National Corpus Sampler. *Language and Computers*, 36:295–306(12).

David Reitter, Johanna D. Moore, and Frank Keller. 2006. Priming of syntactic rules in task-oriented dialogue and spontaneous conversation. In *Proceedings of the Conference of the Cognitive Science Society*.

David Reitter, Frank Keller, and Johanna D. Moore. 2011. A Computational Cognitive Model of Syntactic Priming. *Cognitive Science*.

Sidney J. Segalowitz and Korri Lane. 2004. Perceptual fluency and lexical access for function versus content words. *Behavioral and Brain Sciences*, 27(02):307–308.

Morgan Sonderegger. 2010. Applications of graph theory to an English rhyming corpus. *Computer Speech & Language*, In Press, Corrected Proof.

Svetlana Stenchikova and Amanda Stent. 2007. Measuring adaptation between dialogs. In *Proc. of the 8th SIGdial Workshop on Discourse and Dialogue*.

Richard L. Street and Howard Giles. 1982. Speech accommodation theory. In *Social cognition and communication*. Sage Publications.

Devi Stuart-Fox and Adnan Moussalli. 2008. Selection for social signalling drives the evolution of chameleon colour change. *PLoS Biol*, 6(1):e25, 01.

Paul J. Taylor and Sally Thomas. 2008. Linguistic style matching and negotiation outcome. *Negotiation and Conflict Management Research*, 1(3):263–281.

Jitendra N. Thakerar, Howard Giles, and Jenny Cheshire. 1982. Psychological and linguistic parameters of speech accommodation theory. In C. Fraser and K.R. Scherer, editors, *Advances in the Social Psychology of Language*. Cambridge.

Rick B. van Baaren, Rob W. Holland, Bregje Steenaert, and Ad van Knippenberg. 2003. Mimicry for money: Behavioral consequences of imitation. *Journal of Experimental Social Psychology*, 39(4):393–398.

Arthur Ward and Diane Litman. 2007. Dialog convergence and learning. In *Artificial Intelligence in Education (AIED)*, pp. 262–269.

Danny Wyatt, Jeff Bilmes, Tanzeem Choudhury, and James A. Kitts. 2008. Towards the automated social analysis of situated speech data. In *Proceedings of Ubicomp*, pp. 168–171.

Patrick Ye and Timothy Baldwin. 2006. Verb sense disambiguation using selectional preferences extracted with a state-of-the-art semantic role labeler. In *Proceedings of the Australasian Language Technology Workshop 2006*, pp. 139–148.