Lexical entrainment without conceptual pacts? Revisiting the matching task

Adrian Bangerter (adrian.bangerter@unine.ch)
Institute of Work and Organizational Psychology, University of Neuchâtel
Emile Argand 11, 2000 Neuchâtel, Switzerland

Eric Mayor (eric.mayor@unine.ch)
Institute of Work and Organizational Psychology, University of Neuchâtel
Emile Argand 11, 2000 Neuchâtel, Switzerland

Abstract

In 3 experiments, we investigated whether a decrease in collaborative effort in the matching task occurs even when cards change over trials. We compared two conditions: a classic condition where cards remained the same for each trial, and a new cards condition where cards changed on each trial. Pairs in the new cards condition decreased collaborative effort (albeit less than classical pairs) and lexical diversity. They also were better able to adapt to novel referring situations than classic pairs. Results expand the phenomena relevant for historical accounts of reference.

Keywords: Matching task; collaborative referring; lexical entrainment; conceptual pacts

Repeated referring, lexical entrainment and conceptual pacts

Studies on collaborative referring using the matching task show how partners who repeatedly discuss the same objects require less and less collaborative effort (words) to complete the task (Clark & Wilkes-Gibbs, 1986). This is due to lexical entrainment, the fact that they come to re-use the same words. Lexical entrainment may reflect the elaboration of conceptual pacts, partner-specific agreements about how to name an object that is part of the common ground of conversational partners (Brennan & Clark, 1996). Such phenomena support historical theories of reference design. But can lexical entrainment occur even if partners cannot develop conceptual pacts about specific objects? In everyday life, people with a shared conversational history might find it easy to refer to novel objects for which they have no conceptual pacts. This might happen via (1) semantic or (2) procedural coordination (Mills, 2013). Semantic coordination may lead people to re-use and extend the precedents they have established. If objects change but come from a common universe (e.g., humanoid tangram figures), partners may re-use the same terms to describe recurrent parts of objects (e.g., heads, arms) even without being able to develop conceptual pacts about specific objects. Or they may extend previously used terms to similar objects (Markman & Makin, 1998; Van der Wege, 2009). Procedural coordination involves elaborating shared routines that facilitate task completion, for example schemes for referring to locations in a grid maze (Garrod & Anderson, 1987).

Both procedural and semantic coordination in conversations may facilitate referring to future objects even if they have not been discussed before. Such a phenomenon may expand the scope of historical theories of reference design (Brennan & Clark, 1996). We explored whether this phenomenon holds by implementing it in a variation of the matching task.

Revisiting the matching task

We investigated whether a decrease in collaborative effort in the matching task occurs even when cards change over trials. We compared two conditions: a *classic* condition where cards remained the same for each trial, and a *new cards* condition where cards changed on each trial. We expected pairs in both conditions to exhibit (1) a decrease in *collaborative effort* (measured by a decrease in words), and (2) a corresponding decrease of lexical entrainment (measured by a decrease in *lexical diversity*). But we expected pairs in the new cards condition to show both phenomena to a lesser degree than pairs in the classic condition. We also expected pairs in the new cards condition to rely less on conceptual pacts in repeated referring than pairs in the classic condition (measured by the rate of *indefinite reference*, which is an indicator that referents are not part of common ground). We conducted three experiments with variations on this basic design.

In Experiment 1, we compared the new cards condition with the classic condition. The purpose of this experiment was to investigate whether pairs in the new cards condition are at all able to reduce collaborative effort and whether they exhibit lexical entrainment. Pairs completed 5 trials with eight-card sets.

In Experiment 2, we replicated Experiment 1 and added a sixth trial where pairs in both conditions received new cards. Thus, pairs in the classic condition were confronted with novel cards for the first time, whereas pairs in the new cards condition simply continued as in Trials 1-5. This design tests whether pairs exposed to a wider variety of objects and repeatedly forced to adapt their communication are more flexible than pairs that have repeatedly coordinated via a limited set of conceptual pacts. We expected classic pairs to perform less well on Trial 6 than new-cards pairs.

In Experiment 3, pairs in both conditions initially performed 4 trials. We then switched the matcher in each pair for a second block of 4 trials. This design allowed testing whether collaborative gains in each condition can be transferred from a pair member to a new pair.

Experiment 1

Method

Participants Forty-four unacquainted native French speaking students participated in 22 pairs, 8 in the classic condition and 14 in the new cards condition (anticipating more variance in this condition, we ran more pairs). Individuals were randomly assigned to director or matcher roles; pairs were randomly assigned to conditions.

Procedure Cards were humanoid tangram figures used in previous studies. A computer application presented sequences of 8 cards to participants on separate screens in 5 trials. Cards were presented on the director's screen in a matrix with 4 X 2 placement slots. Matchers viewed an identical, but empty matrix. Their cards appeared below the matrix. They clicked and dragged cards to place them in the slots. In the classic condition, for each pair, a set of 8 cards was randomly drawn from a collection of 40 cards. It remained the same over all five trials. In the new cards condition, pairs placed 5 sets of 8 cards, one for each trial and randomly drawn from the same collection of 40 cards. Dialogue was audio-recorded and transcribed.

Coding We coded collaborative effort, lexical diversity and indefinite reference at the trial level. Interrater agreement was computed for lexical diversity because this measure involved subjective judgments.

Collaborative effort was the total number of words uttered by both participants. Lexical diversity was a ratio of the total of new word types (i.e., not used in previous trials) divided by the total of word types. We focused on content words (nouns and verbs). For each trial, we produced a list containing each uttered noun and verb token. We then combined these into types. We then counted how many of these types had not been uttered in previous trials and divided this number by the total number of types for the trial. Table 1 shows an example. Word types in each column have been uttered during the corresponding trial at least once. Word types in italics are new for the trial in question. Types in plain font have been uttered in a previous trial. In Trial 1, there are 4 types. By definition, none have appeared before, so all are new. Lexical diversity therefore equals 1. In Trial 2, there are four types, but only 2 are new (Square and Diamond) so lexical diversity equals 0.5 (2/4). In Trial 3, there are 3 types, and only one is new (Scarf), lexical diversity thus equals 0.33. In Trial 4, there are 3 types, but none is new. Lexical diversity thus equals 0. In Trial 5, there are 2 types, and again none is new. Lexical diversity thus equals 0. We assessed interrater agreement of the total number of word tokens and of the number of tokens not previously used, as coded by two raters on a subset of the data. Interrater agreement was acceptable (tokens: $r = .75$; tokens not previously used: $r = .72$). Indefinite reference was coded via an automatic content analysis software (Yoshicoder) which counted indefinite pronouns (in French:

un, *une* and *des*) in each trial. For each trial, we divided the number of pronouns by the total number of words.

Table 1: Example coding of word types to compute lexical diversity

Trial 1	Trial 2	Trial 3	Trial 4	Trial 5
Triangle		Triangle		
Head	Head		Head	Head
	Square	Square		
Foot	Foot		Foot	
	Diamond			Diamond
Hand			Hand	
		Scarf		

Results

We analyzed our data using repeated-measures ANOVAs and *t*-tests. Trial was entered as a within-subjects factor, and Condition as a between-subjects factor. Results are displayed in Figure 1.

Collaborative effort decreased over trials and was lower in the classic condition. Because the data violate sphericity assumptions, we report results with Greenhouse-Geisser correction for degrees of freedom. There was a main effect of Trial, $F(2.51, 50.22) = 8.37$, $p < .001$ (linear trend: $p < .001$). There also was a main effect of condition, $F(1,20) = 19.55$, $p < .001$. The interaction between trial and condition was not significant, $F(2.51, 50.22) = 2.23$, $p = .100$.

Lexical diversity decreased over trials. It was lower in the classic condition. ANOVAs were computed between Trials 2-5 because values at Trial 1 equal 1 by definition. We analyzed the difference between Trials 1 and 2 via *t*-tests. Between Trials 1 and 2, lexical diversity decreased in both conditions, classic: $t(13) = 10.23$, $p < .001$, new cards: $t(7) = 18.09$, $p < .001$. Between Trials 2 and 4, there was a main effect of trial, $F(3, 60) = 11.64$, $p < .001$ (linear trend: $p < .001$). There was a main effect of condition, $F(1, 20) = 151.87$, $p < .001$. The interaction between trial and condition was not significant, $F(3,60) = .59$, $p = .63$.

Indefinite reference decreased in the classic condition, but increased in the new cards condition. Because the data violate sphericity assumptions, we report results with the Greenhouse-Geisser correction for degrees of freedom. There was no effect of trial, $F(2.82, 56.53) = 1.73$, $p = .174$. There was an effect of condition: $F(1,20) = 31.37$, $p < .001$. There was an interaction between trial and condition, $F(2.86, 56.29) = 7.96$, $p < .001$.

Collaborative effort

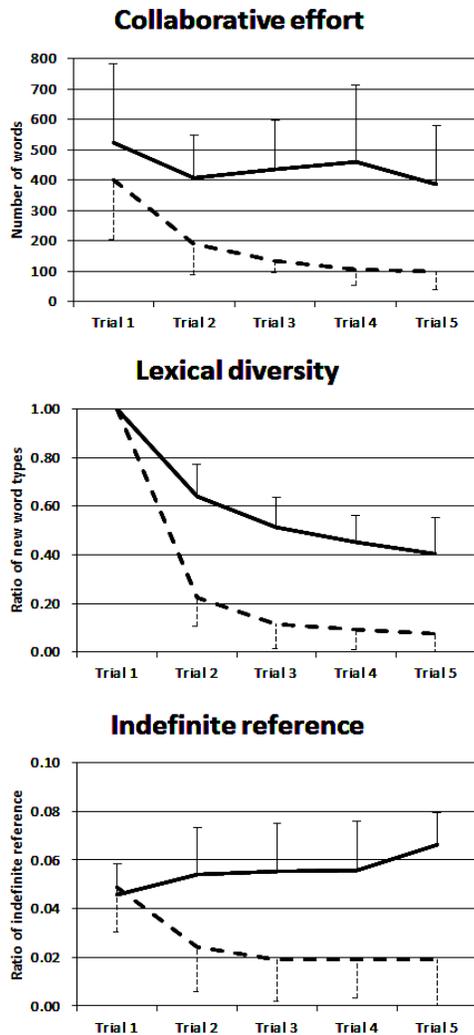## Lexical diversity

Indefinite reference

Figure 1: Collaborative effort, lexical diversity and indefinite reference by trial and condition for Experiment 1. Dashed lines represent the classic condition, plain lines represent the new cards condition.

Discussion

Pairs in the new cards condition reduced collaborative effort over trials, but less than classic pairs. They also reduced lexical diversity, but less than classic pairs. Moreover, they did so without creating conceptual pacts, because indefinite reference ratios remained high in the new cards condition but not in the classic condition. These results suggest that repeated referring to the same objects is not a necessary condition for lexical entrainment. Referring to different but similar objects can facilitate lexical entrainment, and, to a lesser degree, reduction of collaborative effort. However, the reduction in collaborative effort was slight, and there was a lot of variation among pairs. We thus sought to replicate this finding.

Experiment 2

The goal of Experiment 2 was to replicate Experiment 1 findings and to compare effects of repeated referring to the same objects (classic condition) versus referring to different but similar objects (new cards condition) on adaptation to new objects. To this end, pairs in each condition completed five trials as in Experiment 1. For a sixth trial, pairs in each condition received new cards. This amounts to more of the same for pairs in the new cards condition, but confronts pairs in the classic condition with a new situation.

Method

Participants Sixty students participated in 30 pairs (15 in each condition). Participants were randomly assigned to matcher and director roles and pairs were randomly assigned to conditions.

Procedure We used similar cards as in Experiment 1. The computer application this time presented 6 trials in the same way as in Experiment 1. The procedure was exactly the same as in Experiment 1 from Trials 1 to 5. On Trial 6, pairs in both conditions received a new set of cards. Dialogue was audio-recorded and transcribed.

Coding The same variables were coded as in Experiment 1. Collaborative effort was coded in the same way as in Experiment 1. For lexical diversity, word types for each trial were coded using automatic content analysis software. Indefinite reference was coded the same way as in Experiment 1.

Results

We analyzed data using repeated-measures ANOVAs, with Trials 1-5 as within-subjects factor and condition between-subjects. We compared differences between conditions on Trial 6 and differences within conditions between Trials 1 and 6 and 5 and 6 via t-tests. Results appear in Figure 2.

Collaborative effort decreased in both conditions, but more so in the classic condition than in the new cards condition. The sphericity assumption was violated. Thus, we report results with the Greenhouse-Geisser correction for degrees of freedom. There was a main effect of trial, $F(2.65, 74.10) = 25.80$, $p < .001$ (linear trend: $p < .001$) and condition, $F(1,28) = 34.64$, $p < .001$. There was a significant interaction between trial and condition: $F(2.65, 74.13) = 3.92$, $p = .015$. The number of words at Trial 6 was not significantly different in the two conditions, $t(28) = 1.07$, $p = 0.294$. The difference between Trials 5 and 6 was significant in the classic condition, $t(14) = -5.75$, $p < .001$, but not in the new cards condition, $t(14) = .47$, $p = .643$. The difference between Trials 1 and 6 was significant in the new cards condition, $t(14) = 2.41$, $p = .031$, but not in the classic condition, $t(14) = .021$, $p = .984$.

Lexical diversity decreased over trials, and was higher in the new cards condition than in the classic condition.

ANOVAs were computed between Trials 2-5 because values at Trial 1 equal 1 by definition. We analyzed the difference between Trials 1 and 2 via t-tests. Between Trials 1 and 2, lexical diversity decreased in both conditions, classic: $t(14) = 15.62$, $p < .001$, new cards: $t(14) = 18.76$, $p < .001$. Between Trials 2 and 5, there was a main effect of trial, $F(3,84) = 30.61$, $p < .001$ (linear trend: $p < .001$), and condition, $F(1,28) = 47.79$, $p < .001$. The interaction between trial and condition was not significant, $F(3,84) = .796$, $p = .500$. At Trial 6, lexical diversity was higher in the classic condition than in the new cards condition, $t(28)=3.66$, $p = 0.001$. The difference between Trials 5 and 6 was significant in the classic condition, $t(14) = -7.51$, $p < .001$, but not in the new cards condition, $t(14) = -1.98$, $p = .068$. The difference between Trials 1 and 6 was significant in the new cards condition, $t(14) = 21.90$, $p < .001$, and the classic condition as well, $t(14) = 14.96$, $p < .001$.

Indefinite reference decreased over Trials 1-5 in the classic condition but not the new cards condition. There was a main effect of trial, $F(5,140) = 14.56$, $p < .001$, and condition, $F(1,28) = 32.10$, $p < 001$. The interaction between trial and condition was significant, $F(5,140) = 13.09$, $p < .001$. The difference between Trials 5 and 6 was significant in the classic condition, $t(14) = -10.809$, $p < .001$; but not the new cards condition, $t(14)=-.58$, $p < .571$. At Trial 6, indefinite reference was not significantly different in the new cards condition than in the classic condition, $t(28) = 1.044$, $p = .306$. The difference between Trials 1 and 6 was significant in the classic condition, $t(14) = -4.06$, $p < .001$, but not the new cards condition, $t(14) = -1.25$, $p = .231$.

Discussion

We replicated the basic findings of Experiment 1: Pairs in the new cards condition reduced collaborative effort over trials, but not as much as classic pairs. Again, they also reduced lexical diversity, albeit not as much as classic pairs. And again, they did so without creating conceptual pacts, because indefinite reference ratios remained high in the new cards condition but not in the classic condition.

Furthermore, we found that classic pairs showed an abrupt increase in collaborative effort, lexical diversity, and indefinite reference when confronted with new cards. They were less flexible in adapting to new objects of reference than pairs who had repeatedly referred to new objects. Indeed, classic pairs' collaborative effort at Trial 6 was as high as at Trial 1, indicating that they lost all benefits of prior collaboration when dealing with the new cards.

Experiment 3

The goal of Experiment 3 is to replicate Experiments 1 and 2 and to compare the effects of repeated referring to the same objects (classic condition) versus referring to different but similar objects (new cards condition) on adaptation to a new partner. Pairs in each condition completed four trials as in Experiment 1. We then switched the matcher in each pair for a new matcher. The new pairs completed four more trials, with either the same cards they had previously discussed (classic condition) or new cards each trial (new cards condition).

Method

Participants 72 students participated in 24 three-person sessions (12 in the new cards condition and 12 in the classic condition). Participants were randomly assigned to matcher and director roles and three-person groups were randomly assigned to conditions.

Procedure During the first phase of each session, the director performed the matching task over 4 trials with the first matcher. During the second phase, the director performed the task again over 4 trials with the second matcher. In the classic condition, cards remained the same

Collaborative effort

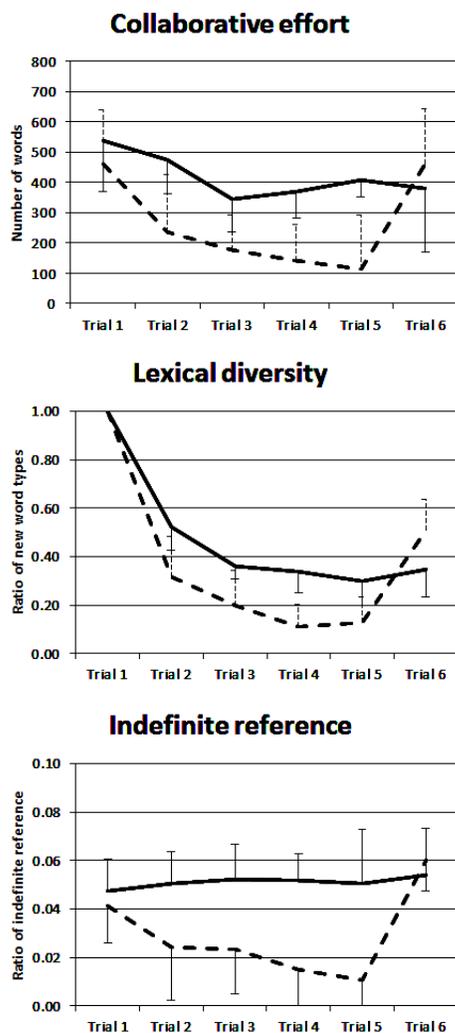**Lexical diversity**

Indefinite reference

Figure 2: Collaborative effort, lexical diversity and indefinite reference by trial and condition for Experiment 2. Dashed lines represent the classic condition, plain lines represent the new cards condition.

over trials and phases. In the new cards condition, cards changed each trial. Dialogue was audio-recorded and transcribed.

Coding The same variables were coded as in Experiment 2.

Results

We ran repeated-measures ANOVAs with trial (the four trials in each phase) and phase (Trials 1-4 versus Trials 5-8) as within-subject factors and condition as between-subjects factor. We compared differences at Trial 5 between conditions and differences between Trials 4 and 5 within conditions via t-tests. We also compared differences between Trials 5 and 1 within conditions via t-tests. Figure 3 displays results.
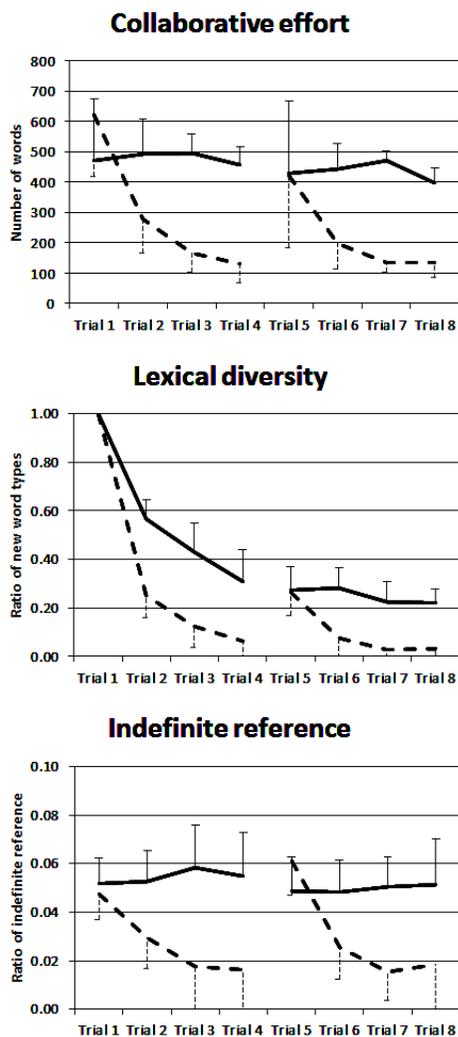

Figure 3: Collaborative effort, lexical diversity and indefinite reference by trial and condition for Experiment 3. Dashed lines represent the classic condition, plain lines represent the new cards condition.

Collaborative effort decreased in the classic condition but not the new cards condition. It also decreased between

Phases 1 and 2. There was a main effect of trial, $F(3,66) = 17.57$, $p < .001$ (linear trend: $p < .001$), a main effect of phase, $F(1, 22) = 10.25$, $p = .004$, and a main effect of condition, $F(1, 22) = 23.53$, $p < .001$. The two-way interaction between trial and condition was significant, $F(3,66) = 18.45$, $p < .001$. The two-way interaction between phase and condition was not significant, $F(1,22) = .87$, $p = .362$. The interaction between trial and phase was not significant, $F(3,66) = 1.38$, $p = .258$. The three-way interaction between trial, phase and condition was not significant, $F(3,22) = 1.54$, $p = .212$. The difference between conditions at Trial 5 was not significant, $t(12) = .12$, $p = .908$. The difference between Trials 4 and 5 was significant in the classic condition, $t(11) = -4.88$, $p < .001$, but not the new cards condition, $t(11) = .42$, $p = .680$. The difference between Trials 1 and 5 was significant in the classic condition, $t(11) = 5.18$, $p < .001$, but not the new cards condition, $t(11) = .94$, $p = .369$.

Lexical diversity decreased more over trials in Phase 1 than Phase 2, and the difference between conditions was higher in Phase 2 than in Phase 1. Again, because lexical diversity equals 1 by definition at Trial 1, the trial variable consisted of the second, third and fourth trials in each phase. We analyzed the difference between Trials 1 and 2 and Trials 5 and 6 via t-tests. Between Trials 1 and 2, lexical diversity decreased in both conditions, classic: $t(11) = 29.27$, $p < .001$, new cards: $t(11) = 18.97$, $p < .001$. Between Trials 5 and 6, lexical diversity decreased in the classic condition: $t(11) = 5.64$, $p < .001$, but not in the new cards condition: $t(11) = -.22$, $p = .832$. For Trials 2-4 and 6-8, there was a main effect of trial, $F(2,44) = 39.75$, $p < .001$ (linear trend: $p < .001$), of phase, $F(1,22) = 106.44$, $p < .001$, and of condition, $F(1,22) = 208.20$, $p < .001$. The two-way interaction between phase and condition was significant, $F(1,22) = 11.10$, $p < .001$, as was the two-way interaction between trial and phase, $F(2,44) = 10.84$, $p < .001$. The two-way interaction between trial and condition was not significant, $F(2,22) = 1.138$, $p = .330$. The three-way interaction between trial, phase and condition was not significant, $F(3,22) = .353$, $p = .704$. The difference between conditions at Trial 5 was not significant, $t(22) = .29$, $p = .775$. The difference between Trials 4 and 5 was significant in the classic condition, $t(11) = -5.57$, $p < .001$, but not in the new cards condition, $t(11) = 1.05$, $p = .316$. The difference between Trials 1 and 5 was significant in the classic condition, $t(11) = 27.64$, $p < .001$, but not in the new cards condition, $t(11) = 25.77$, $p < .001$.

Indefinite reference decreased in the classic condition but not in the new cards condition. There was a main effect of trial, $F(3,66) = 18.86$, $p < .001$. The main effect of phase was not significant, $F(1,22) = .292$, $p = .595$. There was a main effect of condition, $F(1,22) = 32.34$, $p < .001$. The two-way interaction between trial and condition was significant, $F(3,66) = 18.86$, $p < .001$. The interaction between phase and condition was not significant, $F(1, 22) = 3.016$, $p = .096$. The interaction between trial and phase was not significant, $F(3, 66) = 1.99$, $p = .124$. The three-way

interaction between trial, phase and condition was not significant, $F(3,22) = 1.122$, $p = .347$. The difference at Trial 5 between conditions was significant, $t(22) = -2.188$, $p = .04$. The difference between Trials 1 and 5 was significant in the classic condition, $t(11) = -3.71$, $p = .003$, but not the new cards condition, $t(11) = .92$, $p = .377$. The difference between Trials 4 and 5 was significant in the classic condition, $t(11) = -8.66$, $p < .001$; but not the new cards condition, $t(11) = .92$, $p = .380$.

Discussion

Pairs in the new cards condition reduced lexical diversity but not indefinite reference, just like in Experiments 1 and 2. But they did not reduce collaborative effort. Thus, the effect of repeated referring to different but similar objects on collaborative effort is less robust than suggested by Experiments 1 and 2.

Switching matchers had a strong effect on classic pairs. They showed increases in collaborative effort, lexical diversity, and indefinite reference. However, these increases were not as large as the effect of new cards on classic pairs in Experiment 2. Indeed, classic pairs in Experiment 3 did not show as much an increase in collaborative effort as in Experiment 2 when confronted with new matchers. Switching matchers had little effect on pairs in the new cards condition, just like in Experiment 2.

General Discussion

In three experiments, repeated referring to different objects had similar effects to repeated referring to the same objects. Pairs in the new cards condition required less and less collaborative effort to complete the task. They showed lexical entrainment, despite relying less on conceptual pacts than classic pairs. Moreover, they adapted better to novel referring conditions (new cards, a new matcher) than classic pairs. Classic pairs were more perturbed by new cards in Experiment 2 than by a new matcher in Experiment 3.

More detailed analyses (beyond the scope of this paper) should investigate the processes that mediate these differences. First, what do pairs in the new cards condition do differently that helps them better adapt to novel objects? One possibility is that they may elaborate conceptual pacts about parts of figures (e.g., heads, arms). Such pacts may reflect more basic-level categorization than the idiosyncratic pacts that classic pairs converge on. Basic-level categorizations may be easier to extend to novel objects.

Second, it is unclear why pairs in the new cards condition did not decrease collaborative effort in Experiment 3. In all experiments, variance was high in the new cards condition, suggesting differences in how successful individual pairs were in reducing collaborative effort. However, lexical diversity decreased and substantially in all experiments. This suggests that differences between individual pairs are not due to a failure to entrain. One possible explanation is that pairs that did not reduce collaborative effort expended more effort into reducing the uncertainty of initial referring expressions. That is, there is a difference between when pairs have an initial hypothesis about the identity of a referent and when they are both certain that they are talking about the same referent, and collaborative effort is necessary to attain this certainty (Brennan, 2005).

Third, we focused on semantic coordination. But pairs in both conditions may also have been able to reduce effort in procedural coordination (Mills, in press). For example, they may have discussed card placement (e.g., the details of the 4 X 2 grid) in less detail over trials. Improvements in procedural coordination have not been analyzed in previous matching task studies, and may also contribute to the overall improvement in collaborative effort beyond the effect of lexical entrainment via conceptual pacts.

Our findings have implications for further investigation of historical factors in referring. Repeated referring to the same objects by conversational partners leads to elaboration of partner-specific conceptual pacts which reduce collaborative effort. But repeated referring to different but similar objects also fosters lexical entrainment and helps reduce collaborative effort. It also enables more flexible adaptation to new referring situations. Repeated referring to similar objects may constitute a more general process by which conversational partners build common ground than repeated referring to the same objects.

References

Brennan, S. E. (2005). How conversation is shaped by visual and spoken evidence. In J. Trueswell & M. Tanenhaus (Eds.), *Approaches to studying world-situated language use: Bridging the language-as-product and language-action traditions* (pp. 95-129). Cambridge, MA: MIT Press.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology, Learning, Memory, and Cognition, 22*, 1482-1493.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition, 22*, 1-39.

Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: a study in conceptual and semantic co-ordination. *Cognition, 27*,181-218.

Markman, A. B., & Makin, V. S. (1998). Referential communication and category acquisition. *Journal of Experimental Psychology: General, 127*, 331-354.

Mills, G. J. (in press) Dialogue in joint activity: complementarity, convergence, conventionalization. *New Ideas in Psychology*.

Van der Wege, M. M. (2009). Lexical entrainment and lexical differentiation in reference phrase choice. *Journal of Memory and Language, 60*, 448-463.