Joint Acquisition of Path and Manner Action Description

Claire S. Shah (shah.179@wright.edu)
Department of Psychology, 335 Fawcett Hall, 3640 Colonel Glenn Hwy.
Dayton, OH 45435

Valerie L. Shalin (valerie.shalin@wright.edu)
Department of Psychology, 335 Fawcett Hall, 3640 Colonel Glenn Hwy.
Dayton, OH 45435

Abstract

The present study examines language patterns in the formation of common ground in collaborative action tasks. Based on the classic Clark and Wilkes-Gibbs' (1986) paradigm for object descriptions, we examined dialogue between pairs of participants as they work cooperatively to maneuver a remote control car following both manner and path instructions. Overall, we replicated Clark and Wilkes-Gibbs' (1986) results in the domain of action in the decline of word count, verb phrases, turn taking, and number of errors committed, with diminishing returns after one trial. However, we also document specific language reductions in path related actions, but not in manner related actions. We suggest that path actions particularly depend on compositional descriptors of the environment, consistent with the contemporary conceptualization of action (Barsalou, 2009).

Keywords: common ground; action description

Introduction

As individuals learn a new collaborative task and attempt to communicate, they settle on a lexicon to describe task elements. Clark and Wilkes-Gibbs (1986) showed that pairs learn to use fewer and more concrete noun phrases in collaborative communication to describe an unfamiliar set of objects in a sequencing task. However, their simple object sequencing task did not require referents to action, which is the foundation of distributed work. Nevertheless, traditional linguistics suggests a generalization of grammar rules across noun and verb phrases (Chomsky, 1970). Therefore, similar patterns should emerge in the establishment of lexical referents for both noun and verb phrases. This research reveals the limitations of this generalization, by distinguishing between the acquisition pattern for lexical referents to manner and path action (Talmy, 2000).

Referential Language

In their seminal paper, Clark and Wilkes-Gibbs (1986) showed that as a pair of individuals work cooperatively to complete a novel-object sequencing task, the amount of language necessary to identify initially unfamiliar task objects decreases. In their study, builders arranged abstract tangram images in an order dictated by a director. In each trial, the order of the same tangram images was randomized, and the participants completed six trials. Clark and Wilkes-Gibbs (1986) recorded the time it took participants to complete the ordering task, and transcribed all spoken communication including "changes of speaker, back-channel responses, parenthetical remarks, interruptions, hesitations, false starts, and basic intonational features" (p. 11). They found significant declines in changes of speaker, time on task, total word count, and noun phrase count. They explained their results using the concept of "common ground" developed over trials. The development of a shared lexicon reflects common ground and reduces communication requirements, in turn reducing time on task. In particular, the number of noun phrases exchanged decreased after each trial.

Long-standing linguistic theory argues that noun and verb phrases are governed by the same underlying grammatical structure (Chomsky, 1970), anticipating the same language acquisition patterns for the sequencing of previously unfamiliar action as objects. We examined the verbal description of actions in a task requiring the sequencing of maneuvers of a remote controlled car.

Principled Selection of Action Stimuli

Cognitively-oriented linguists (Talmy, 2000) make a distinction between types of action, and how this distinction is addressed in different languages. In English, specific verbs identify the *manner* in which an action is conducted. For example, separate words distinguish walking from running and skipping. Studies have shown similarities in the encoding of motion words across multiple languages (Papafragou & Selimis, 2010), but here we focus on English.

With rare exceptions such as enter and exit, expression of *path* in English requires a preposition, e.g., walk into/by/next to, etc. We note additionally that prepositions require arguments, that is grammatical objects. We walk into *the house* or by *the lake*. Path referents are necessarily compositional; manner referents are not. As a result, substituting action for object in the Clark and Wilkes-Gibb sequencing task requires sensitivity to the type of action in question, as the potential for truncated referents over trials is not the same. Moreover, empirical evidence regarding the acquisition of truncated references that distinguishes between action type questions the psychological relevance

of the purely grammatical equivalence of noun phrases and verb phrases.

Several linguists have attempted a semantic categorization of action (Jackendoff, 1991; Talmy, 2000). Schank (1972) identified 11 conceptual primitives: ATRANS, ATTEND, INGEST, EXPEL, GRASP, MBUILD, MTRANS, MOVE, PROPEL, PTRANS, SPEAK. For this study, we sampled from PROPEL (apply force to an object) for manner related actions and PTRANS (change physical location of an object) for path related actions. These conceptual primitives aided in the classification and distinction between path, manner, and combination maneuvers specified in the following section. PROPEL examples of manner may be more amenable to single word capture while PTRANS examples may require multi-word noun phrases due to the persisting need to articulate location. On the other hand, PTRANS examples might be particularly amenable to truncated terminology as participants gain familiarity with the objects and their relationships in a constant task environment. This coupling of action with the environment is consistent with contemporary views on the mental representation of action (Barsalou, 2009).

Linguistic Characteristics

Since the original Clark and Wilkes-Gibb experiment, new methods have emerged to conduct automated content analysis on unstructured text, allowing us to enrich simple measures of word count easily. In lieu of tallying specific lexical items, Pennebaker and colleagues' (2015) Linguistic Inquiry and Word Count (LIWC) software enables the automatic tallying of abstract categories of words, generally as a proportion to control for text length. Of the approximately 90 pre-defined LIWC categories, we chose *a priori* 19 to focus our analysis, including word count and common verbs for parity with the original Clark and Wilkes-Gibbs' (1986) study. To ensure all verb phrases were counted, the categories for auxiliary verbs and common adverbs were analyzed. The other parts of speech we chose to analyze are prepositions, interrogatives, and comparatives. The overarching cognitive processes category was chosen to identify the underlying thought processes of the chosen language. This category consists of insight words, causations, discrepancies, tentativeness, and certainties (Pennebaker et. al., 2015) perhaps best associated with metacognitive processes, or backchannel functions in dialogue specifically measured by assent words. We hypothesized reductions in this content across trials as participants established common ground. The overarching time orientation category was chosen to further investigate the action orientation of the stimuli. This includes motion, space, time, past focused words, present focused words, and future focused words. The later analyses were purely exploratory without any hypotheses. Having obtained the pattern of results for the *a priori* measures, we also examined adjectives as a proxy for noun phrases.

Method

Participants

Fifty undergraduate-level students (25 pairs) participated in this study in exchange for laboratory participant credit in an introductory psychology class. Data from fourteen pairs were discarded due to extenuating factors including not meeting English requirements as stated in the recruitment requirements, equipment failures, and failure to complete all six trials (to exclude incomplete data). Data from the remaining 11 pairs are used in this study. This is an increase in participants compared to Clark and Wilkes-Gibbs (1986) study, which included data from 8 pairs. Participants included 59% females and 41% males with an average age of 22.41 years old (M = 20.76 removing a 59 year old outlier).

Equipment

We used a Tera WLtoys A999 1:24 Electric 2WD Remote Control RC car. A set of car maneuver videos were pre-recorded in the experiment room, from the perspective of the driver, and displayed on a first-generation iPad Air. The maneuver sequences for trials two through five were randomized. The first and sixth (last) maneuver sequences were consistent throughout all participant pairs to increase comparability across participant pairs.

Stimuli

The six chosen maneuvers are divided into categories based on verb type. Below is the list of each maneuver with their description and predetermined error criteria. These six maneuvers were chosen out of 16 pilot tested maneuvers to limit participant fatigue. Though categorized by path, manner, or combination for design purposes, these categories were not identified for the participants.

Figure 1: Remote control car in the experimental set up

List of Maneuvers and Corresponding Error Allowances

Path Maneuvers

1. Place the front right tire onto an object on the floor (paper circle) – must place only the front right tire on the paper circle. It is incorrect if the driver places any other part on the circle.
2. Drive the car in 1.5 circles – Must end at the same spot as the car in the video (perpendicular to the single cone facing west) within 2 inches

Manner Maneuvers

1. Move the car back and forth 8 times in short bursts – short bursts are about 6 inches in length
2. Lightly tap the front of the car on the back mirrored wall – must approach the wall slowly, tap can not make a sound

Path and Manner Combined Maneuvers

1. Reverse in a straight line against the ramp – must begin at the top right corner of the ramp and end on the bottom right corner.
2. Reverse into a "parking space" – must reverse into the cones and follow the same pattern as the video.

Task

Participants were randomly assigned to either the role of director or driver. The director watched un-narrated video footage of the remote-control car executing each maneuver in series and verbally directed the driver on how to execute each one correctly. The driver used the director's verbal instruction to execute the sequence of maneuvers correctly. The driver and the director could not see each other, and the director could not see the experimental environment. Once the participants had completed a maneuver, the experimenter either stated "correct" or "incorrect". After each maneuver, the car was returned to the center of the room on a white X. This allowed independent analysis of the language used for each maneuver. If the experimenter stated "correct", then the participants moved onto the next maneuver. If "incorrect", they repeated the same maneuver.

Measures

Process. All trials were video and audio recorded. The first author transcribed each pair's conversation into Word documents. The conversations between the director and driver (omitting experimenter intervention) were then run through Language Inquiry and Word Count (LIWC) software to determine the linguistic characteristics of the dialogue (Pennebaker et al. 2015). Word count was measured using the LIWC software as an integer. LIWC produces an output of *proportions* of the different content types. Proportions enabled content analysis independent of total word count, as we predicted total word count would decrease over trials. Specific LIWC categories of interest were prepositions, auxiliary verbs, adverbs, compare, insight, discrepancy, verbs, tentative, certain, focus on past, focus on present, focus on future, motion, space, time, assent, and interrogatives. Verb phrases were identified by the proportion of verbs given in the output from LIWC. Proportion of adjectives served as a proxy for noun phrases.

Outcome. Turn-taking was defined as the number of times the speaker switched between the driver and director. Any communication exchanges with the experimenter were excluded from this measure. Errors for each maneuver were determined during the experiment using the corresponding error allowances described in the above task section. When a participant pair performed a maneuver outside of the error allowance, the experimenter stated "Incorrect" which informed the pair that they needed to restart that maneuver and make another attempt. During analysis, the number of errors was determined by the number of times the participant pair had to restart a maneuver within a trial.

Procedure

The participants assigned to the role of driver had 15 minutes to familiarize themselves with the controls of the car. Formal training was not conducted to avoid any potential bias towards using specific words used by the experimenter. Next, the director watched a video of 6 different clips of maneuvers executed by the car. After watching the clips, the director instructed the driver on how to execute each maneuver correctly. Once the pair successfully completed the first trial of all six maneuvers, the driver and the director repeated the same procedure five more times. Each new trial had the same maneuver clips, but in different orders. The first and last trial were set as described above and trials two through five were randomized for each participant pair. The experiment concluded once the participants completed all six trials or when the allotted two hours elapsed, whichever came first. The pairs that did not finish by two hours were not included in the data analysis.

Results

General Trial Effects

All analyses were conducted as a repeated measure ANOVA, with trial as a nominal predictor. We report significance using Huynh-Feldt critical values, and follow-up with linear and quadratic trend analysis. Consistent with previous work, overall declines in word counts across the trials as well as overall declines in verb phrase count were observed. An anticipated positive trend in accuracy and an anticipated decline in the number of turns (alternations in speaker) taken by each participant in the pair were also found.

Verb Phrase Count. There was a decline in the proportion of verbs overall as participants repeated the six-maneuver sequence for six trials, $F(4.30, 43.02) = 4.36$, $p < .01$, $\eta^2 = .11$, with a significant linear trend ($F(1,10) = 8.89$, $p < .05$).

The example below illustrates the difference between the verb (*verb phrases italicized*) usage from trial one to trial six for the same pair for a path maneuver as well as key environmental referents (**environmental referents are bolded**). Note initial, dismissed concern for manner ("*how fast*"), the persisting reference to the environment ("**the white circle**"), and the emergence of a single word referent for the maneuver ("J").

Path Maneuver #1
Trial 1 (segment)

Director: *You're gonna go start going straight* but *don't go past* the basketball or any cones and then kinda *go in a curve*, like a right curve. *So you're just gonna go like* inside *think of the basketball* and the cones
Driver: Ok
Director: As like a border and *stay inside* the border and end at the white square
Driver; Alright so uh *does it does it* matter how fast *I do it*?
Director: Just like a normal like a constant constant speed I *guess*. Kind of fast. And then end at like the edge of the **white circle**, outer edge towards the cones
Driver: Ok. *Do I* stop before the **circle** *starts?*
Director: Um, on the **circle**, on the outer edge of the **circle**
Driver: Alright
(Successfully completed maneuver)

Trial 6
Director: The next one *is the J* and you end on the **white** um **circle**
(Successfully completed maneuver)

Word Count. We identified a statistically significant reduction in the mean number of total words used by participants across trials (F (1.10, 11.02) = 31.10, $p < .001$, $\eta^2 = .69$). (See Figure 2). There was support for both linear and quadratic trends ($F(1,10) = 34.43$, $p < .001$ and $F(1,10) = 29.97$ $p < .001$ respectively).

Task Accuracy. There was a statistically significant decline in the number of errors as the participants progressed through the trials (($F(1.27, 12.68) = 38.28$, , $p < .001$, $\eta^2 = .77$). There was support for both linear and quadratic trends ($F(1,10) = 35.13$, $p < .001$ and $F(1,10) = 46.89$ $p < .001$ respectively).

Task accuracy was measured by the number of errors committed by the participant pairs. Number of errors was defined as the number of incorrect attempts made by the pairs. Figure 3 illustrates the number of errors by participant pairs for each trial. Errors declined most from trial one to trial two, and no pairs had more than two errors from trial two onwards.

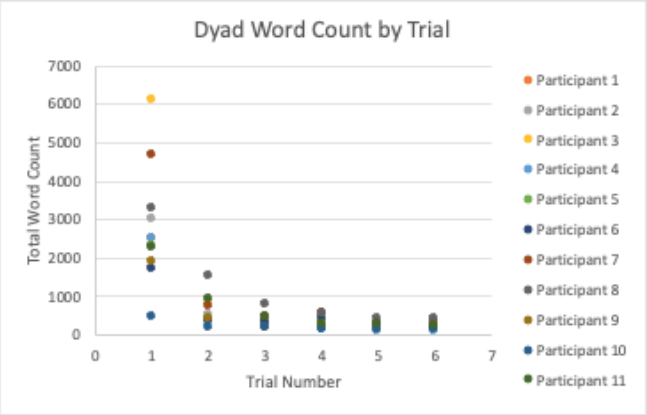

Figure 2. Word count per trial per participant pair

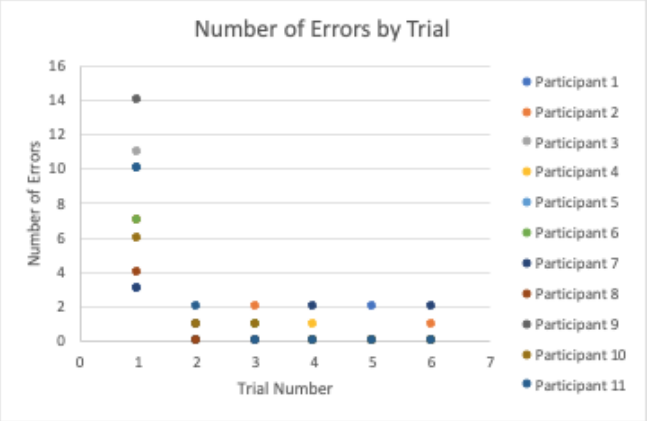

Figure 3. Number of errors per trial per participant pair

Turn-taking. As the participant pairs progressed through the trials, the number of conversational turns decreased ($F(1.09, 10.86) = 25.28$, , $p < .001$, $\eta^2 = .62$). There was support for both linear and quadratic trends($F(1,10) = 28.58$, $p < .001$ and $F(1,10) = 23.60$ $p < .01$ respectively). Figure 4 illustrates the substantial decline from trial 1 to trial 2.

One striking example of decline occurred between the driver and director in participant pair 3. In the first trial, they alternated turns 387 times before successfully completing the trial. By the last trial, they only alternated turns 12 times - two per maneuver where the driver's responses were all agreement words to acknowledge that he/she understood which maneuver to complete.

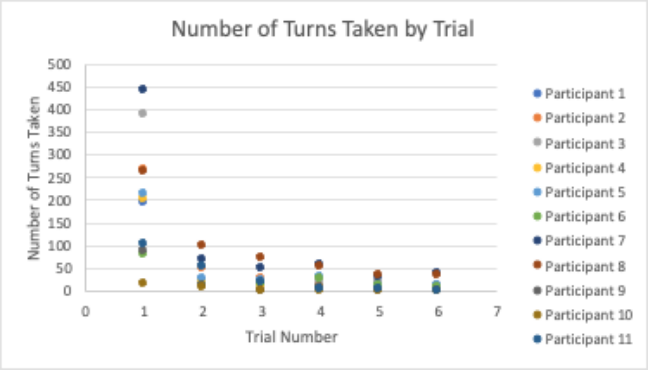

Figure 4. Number of conversational turns per trial per participant pair

Below is an example of the dialogue between a director and driver for the same manner maneuver (*manner related terms are italicized*) in trial one and then again in trial six (**environmental referents are bolded**). The example demonstrates that the number of turns taken decreased from trial one to trial six, but also shows the phrase the two participants had come to understand to initiate that maneuver. In trial one, we see the manner words "*ever so slightly"* and the environmental referent "**bottom of the mirror**" that persist in trial six.

Manner Maneuver #2
Trial 1:
Director: You're gonna start driving towards the mirror *ever so slightly* go to the left and then straighten back out, so you go right in the middle between the basketball and closest cone to it's right
Driver: Ok
Director: And then just go all the way forward until the wheels touch the **bottom of the mirror,** and then stop there
(Driver successfully completed maneuver)

Trial 6:
Director: The first maneuver is the same one we just did where you go up to **the bottom of the mirror** *ever so slightly*
Driver: Not too hard
(Driver successfully completed maneuver)

Maneuver Type Effects
The linguistic categories of interest were analyzed using LIWC. The output consisted of the *proportion of each category* within each trial to ensure independence from the decline in total word count. Tables 1-4 illustrate any linguistic measures that had significant linear or nonlinear trends in the whole dialogue and by maneuver type. All maneuvers display decreases in word count.

We attempted to explain the reduction in word count with LIWC proportion measures. Maneuver type differed in the initial number of words used for manner (M = 648, SD = 618), path (M = 900, SD = 642), and combination (M = 1214, SD = 618) maneuvers, with less final reduction in word count for manner maneuvers (.12) than path or combination maneuvers (.06). Separate analyses for each maneuver type, path and combination type maneuvers attempt to explain these as content reductions.

Table 1. Summary table of significant results (F-values) examining trial effect by maneuver type.

Measures	Whole Dialogue	Path	Manner	Combination
Word Count	31.10***	16.69**	11.51**	16.54**
Focus on Past	3.83*	2.49*	2.63*	
Discrepancies	3.77*	3.14*		3.56*
Verbs	4.36**			3.36*
Tentativeness	4.96*	3.56*		3.24*
Prepositions		2.74*		
Certainty	3.01*			
Causalities	4.94**	4.61**		
Focus on Present	3.85**			3.74*
Interrogatives		2.55*		
Auxiliary Verbs	3.11*			

Note: * = $p < .05$, ** = $p < .01$, *** = $p < .001$ using Huynh-Feldt critical values.

The preponderance of significant specific changes occur for path and combination maneuvers. Most of the changes are reductions in the proportion of a language feature. Participants use less tentative language and fewer verbs and call out fewer discrepancies. No such changes (and no other examined potential changes) were observed for manner maneuvers, which, because of their initial smaller word counts, had potentially less room to contract.

Observing this striking difference in acquisition patterns for manner maneuvers, we examined a proxy for noun phrases (adjectives) under the suspicion that these would persist in path maneuvers due to their emphasis on the environment. In fact, path maneuvers displayed a significant positive pattern between adjectives and trial number. As shown in the example below *(adjectives are italicized)*, there are more adjectives in trial one than trial six, but the proportion of adjectives increases in the final trial because the exchange is much shorter. This confirms the importance of the environment in the characterization of path maneuvers.

Path Maneuver #1
Trial 1
Director: Um, starting from the X, you make a right turn and then drive towards the *white circle*...Ok, um, drive up towards the basketball, make a right there,

follow along the triangles down and then end on the *white circle*
Driver: Ok. Do I need to weave in between the cones?
Director: You don't go in between you kind of just, starting at the X, you're coming forward, it's almost like it's drawing a nine, so like if
Driver: Should I go around the left of the basketball?
Director: The right of the basketball
Driver; Ok, right of the basketball
Director: So yea, you're turning right at the basketball. It's a pretty fast turn, it's a fast right turn and um, and end at the
Driver: Then go down all the cones, go around the last cone and stop at the *white circle*
Director: Uh, don't go around any of the cones, you should be in between the X and the cones in the little space between
Driver: Ok
Director: Um, are you, *as soon as you* leave the x, drive a little left towards the *top corner* of the, um, uh the *long orange thing*, I don't know what it is, a tunnel? You kind of go close to that corner and then make your right turn. And that should be *clear*
Driver: Um, so
Director: Ok, um, start at the X, make a *small curve* towards the left, once you the car is at the corner of the tunnel, you start driving a little towards the basketball and immediately make a sharp right and then follow all the way down to the *white circle* without going through any cones
Driver: Now when I take that sharp right, do I go around this first cone?
Director: You don't go around any cones, you kind of like, um, yea, you don't go around the cones
Driver: Ok
(Driver successfully completed maneuver)

Trial 6
Director: Ok, for the *next one* you're doing the u turn at the basketball and you're kinda drawing that nine and end at the *white circle*.
(Driver successfully completed maneuver)

Discussion and Conclusions

We extended the classic Clark & Wilkes-Gibbs paradigm in the space of objects to a task in the space of actions. We found many results in these action-based tasks that are consistent with Clark and Wilkes-Gibbs. Total word count and verb phrase usage significantly decreased as participants progressed through the experiment. The number of turns each participant took to speak also decreased significantly over time. The number of errors committed declined over trials. As in Clark and Wilkes-Gibbs, participants communicated more effectively over time after they achieved common ground and created new agreed upon terms for maneuvers.

Further we found that our ability to explain the word count reduction depended on maneuver type, suggesting that something different is occuring in the process of establishing common ground for at least some types of action that is not relevant to the establishment of terminology in the space of objects. Path related action is associated with a reduction in the proportion of several LIWC measures, and a likely increase in the proportion noun phrases suggested by the increase in adjectives. This emphasis on the environment is consistent with the idea that the mental representation of action is closely coupled with the objects it involves (Barsalou, 2009).

However, manner related maneuvers show none of these effects, despite an overall reduction in word count and focus on past related words. In particular, we do not see a reduction in the proportion of verbs for manner maneuvers. By illustrating a difference in the acquisition of manner and path descriptions, we reinforce a psychological difference between noun phrases and at least some verb phrases despite their linguistic generalizability. This is not to argue that participants fail to reach common ground in the description of manner action, but that it is not accomplished in the same way as path maneuvers, which depend on familiarity with the objects in the environment. This also supports a conceptual distinction, consistent with the idiosyncrasies of English, between path and manner verbs. Nevertheless, we cannot rule out that this distinction may only be present in verb-framed languages such as English and we could be observing an artifact of the language.

However, maneuver type may have influenced the pattern of findings for unexpected reasons. Zacks' (2020) analysis of events focuses on the consistency with which observers segment them. Intermediate stationing at the center of the room and a randomized ordering of maneuvers established maneuvers as the unit of analysis. Moreover, our manner maneuvers do not have multiple segments to either truncate or organize as a higher order named unit. Such an explanation remains consistent with a distinction between manner and path, the language that refers to object versus action and perhaps between the conceptualization of objects and actions themselves.

The dependence of path action on the environment raises concern for transfer to other environments. We have demonstrated that the environment is influencing and directing the language used in forming common ground between the director and driver. If we change the environment, we expect to see similar patterns emerge while learning the new referents needed to complete the task successfully—little or no carryover. The observed pattern of findings in our experiment also raises the question of how mostly environment-independent manner terminology is acquired. We speculate that manner terminology is

culturally elevated with designated, non-compositional lexical items and particularly dependent on learning by being told.

References

Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1281-1289.

Chomsky, N. (1970). Remarks on nominalization. In: R. Jacobs and P. Rosenbaum (eds.) Reading in English Transformational Grammar. Waltham: Ginn.

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

Jackendoff R., (1991). Parts and boundaries. *Cognition, 41,* 9-45.

Papafragou A., & Selimis S. (2010). Event categorisation and language: A cross-linguistic study of motion, *Language and Cognitive Processes*, *25*(2), 224-260.

Pennebaker, J.W., Booth, R.J., Boyd, R.L., & Francis, M.E. (2015). Linguistic Inquiry and Word Count: LIWC2015. Austin, TX: Pennebaker Conglomerates (www.LIWC.net).

Schank, R. (1972). Conceptual dependency: A theory of natural language understanding. *Cognitive Psychology*, *3*, 552-631.

Talmy, L. (2000). *Toward a cognitive semantics*. Volume I: Concept structuring systems. Cambridge, MA: MIT Press.

Zacks, J. M. (2020). Event perception and memory. *Annual Review of Psychology*, *71*, 165-191.