

Beyond canonical form: Verb-frame frequency affects verb production and comprehension

Lise Menn,^a Susanne Gahl,^b Audrey L. Holland,^c
Gail Ramsberger,^a and Daniel S. Jurafsky^a

^aUniversity of Colorado

^bUniversity of Illinois

^cUniversity of Arizona

A language's canonical form is considered its easiest clause structure (major constituent order) to process. This finding rests primarily on comparing active and passive frames of transitive verbs like 'kiss.' In the active voice, the subjects of these verbs are agents, while the objects are undergoers/items affected by the action. However, recent studies suggest that the unaccusative verb frame (*the apple dropped*)—roughly, intransitive with surface subject playing the same role as the surface object of a corresponding transitive (*someone dropped the apple*)—is no more difficult than the intransitive with agent subject (Gottfried, Menn, & Holland, 1997, using repetition; Gahl, Menn, Ramsberger, Jurafsky, Elder, & Rewega et al., 2001, using plausibility judgement). Meanwhile, transformational accounts of aphasic syntax that ascribe the difficulty of the passive voice (*the apple was dropped*) to problems in computing traces left by movement transformations predict that unaccusatives should be difficult for people with agrammatic aphasia in the same way as passives (Kegl, 1995), since they are derived from underlying strings by similar movement rules.

Analyses of large language corpora, made possible by computational linguistics, have established the frequency of use of various argument structures for many English verbs (Gahl, Jurafsky, & Roland, 2003). Many verbs have a statistically preferred 'favorite' frame: *deny* is **P-bias**, i.e., most often used in the Passive frame (*the report was denied*); *lean* is **IA-bias**, i.e., prefers an Intransitive Active voice frame with the subject as agent (*she leaned against the wall*); and *spill* is **IU-bias**, i.e., prefers the Intransitive Unaccusative (*the milk spilled on the floor*). Gahl (2002) showed that such lexical frequency biases might underlie parsing preferences for a variety of transitive and intransitive verbs in normal and aphasic speakers.

Gahl et al. (2001) used a plausibility judgement task to show that a mixed group of aphasics performed significantly better on unaccusatives than on passives, and that the difficulty of an argument structure was modulated by lexical bias. Sentences whose structure matched the lexical bias of the main verb (concordant condition) were significantly easier than sentences in which structure and lexical bias did not match (discordant condition).

We have extended the judgment results of Gahl et al. (2001), and obtained production data as well. S.K., a woman with Broca's aphasia, judged 264 sentences (half plausible, half implausible) and repeated 98 plausible sentences. Verb frame biases were computed from the TASA corpus of school reading materials. Thirty verbs were used: 6 IA-bias verbs that can also be used as transitives (*sail*), 6 IU-bias that can also

be used as transitives (*spill*), 6 P-bias that can also be used as active transitives (*clean*), and 12 Transitive (T)-bias verbs: 6 whose subjects keep their semantic role when intransitive (T (IA), *we ate supper/we ate*) and 6 whose transitive objects become their subjects (T (IU), *she shattered her glasses/her glasses shattered*). Each verb was used with two different sets of concrete nouns, and each verb/noun set combination used in at least two frames, its favorite frame (concordant condition) plus another frame (discordant condition): IA-bias verbs in T frames, IU-bias in T frames, P-bias in T frames, and T (IU) in IU frames. The T (IA) verbs were presented in three frames: T, IA, and P, yielding 11 combinations of type and frame. Each judgment item was presented in a plausible and an implausible variant (*The room was cleaned by the maid this morning; The maid was cleaned by the room this morning*), yielding 24 sentences for each combination of verb type and verb frame.

Judgment results are in accordance with our predictions except for the passives. IU (unaccusative) frame judgments overall were better than passive and equal to IA; indeed, performance on IU-bias verbs in IU frames was better than any other cell (22/24 correct). For T-, IA-, and IU-bias verbs, performance was best in the concordant frame. However, contrary to prediction, T-bias verbs in Passive frames were judged better than P-bias verbs in Passive frames.

Repetition stimuli were versions of 98 of the plausible judgment sentences, which were shortened but still too difficult for SK to repeat verbatim. As predicted, verbs in their favorite (concordant) frames were more easily processed than verbs in discordant frames: they yielded better-structured responses (52% vs. 32% scorable argument structures) and the modeled structures were preserved more frequently (83.3% vs. 60%). Changes to the input structure which yielded scorable responses tended to replace discordant frames by concordant ones, but not the reverse, unless the model or the favorite frame was passive (see Table 1).

Our findings challenge a transformational account of aphasic speakers' problems with verb argument structures. They instead support an account that considers syntax, semantics, and construction frequency, sharing many of the viewpoints of Construction Grammar (Goldberg, 1999) and computational psycholinguistics. Aphasic speakers' apparent dependence on Agent-first canonical form may be dependence on active voice verb morphology, plus an expectation that the mapping of noun positions onto semantics will follow the pattern of the individual verb's most frequent argument structure pattern (cf. Menn, 2000).

Table 1

Repetition results for 98 sentences: Stimuli in concordant frames are more scorable and more likely to maintain structure when repeated, except for passives

Model sentence frame	Verb type												Total
	IA-bias		IU-bias		T-bias (IU)		T-bias (IA)			P-bias			
	IA	T	IU	T	T	IU	T	IA	P	P	T		
Number of model sentences in frame	10	9	8	9	14	10	5	8	7	9	9	98	
Number of scorable sentences in frame	5	3	3	3	11	6	3	4	2	2	2	44	
Number of scorable sentences in frame that maintain model structure	4	1	3	1	10	5	3	3	1	0	1	32	
Number of scorable sentences in frame that change structure to favorite	n/a	2	n/a	2	n/a	1	n/a	1	0	n/a	0	6	
Number of scorable sentences that change structure to non-favorite	1	0	0	0	10	0	0	1>	IU2>	T1>	*IA	6	

References

- Gahl, S. (2002). The role of lexical biases in aphasic sentence comprehension. *Aphasiology*, 16, 1173–1198.
- Gahl, S., Menn, L., Ramsberger, G., Jurafsky, D.S., Elder, E., & Rewega, M., et al. (2001). Syntactic frame and verb bias in aphasia: Plausibility judgments of undergoer-subject sentences. Twelfth Conference on Theoretical and Experimental Neuropsychology, Montreal 2001. Scheduled for publication in *Brain and Cognition*, 2003.
- Gahl, S., Jurafsky, D.S., & Roland, D. (2003) (ms. submitted). Verb subcategorization frequencies: American English corpus data, methodological studies, and cross-corpus comparisons.
- Goldberg, A. (1999). The emergence of the semantics of argument structure constructions. In: Brian MacWhinney (Ed.), *The Emergence of Language*. Mahwah NJ: Lawrence Erlbaum Associates.
- Gottfried, M., Menn, L., & Holland, A.L. (1997). Verb-argument structure and aphasic repetition. Poster, 1997 meeting of the Academy of Aphasia, Philadelphia.
- Kegl, J. (1995). Levels of Representation and Units of Access Relevant to Agrammatism. *Brain and Language*, 50, 151–200.
- Menn, L. (2000). It's time to face a simple question: What makes canonical form simple? *Brain and Language*, 71, 157–159.