

# Why Subject Relative-Clauses Prevail: A Corpus Study

Laura Michaelis and Jill Duffield  
Linguistics, CU

Proponents of Construction Grammar (e.g., Goldberg 2006, Sag forthcoming) tend to attribute linguistic facts to idiomatic grammatical patterns rather than general-purpose processing or encoding constraints. Construction Grammar therefore strikes its critics as descriptive rather than explanatory. In this talk, we will weigh competing modes of explanation for a well-known structural preference of English, and demonstrate that a construction-based explanation is superior. The usage trend at issue involves subject relatives, relative clauses that contain a relative pronoun where a subject noun phrase would go:

- (1) I've got a friend [that uh lives right next to the Cowboys' training camp].
- (2) There's a lot of people [who fall into that category].

Subject relatives contrast with those relative clauses that contain a gap in place of a postverbal argument:

- (3) I always like the letters [that they read\_\_].

Subject relatives are the prevalent type both across languages (Keenan and Comrie 1977), and in conversational speech, accounting for 65% of relative clauses in the American National Corpus (Reali & Christiansen 2007) and 67% in the corpus under study here, the Switchboard corpus. This fact could be attributed to processing constraints, as per Hawkins 1999, 2004: subject extractions are the most local filler-gap dependency. But the processing account fails to explain a widely observed bias in English corpora: subject relatives are preferred only as modifiers of objects, e.g., *a friend* and *a lot of people* in (1-2) above; they are

rare as modifiers of subjects (Fox & Thompson 1990, Geisler 1998). Further, recent findings cast doubt on the putative processing advantage for subject relatives (Mak et al. 2008). We offer an alternative, construction-based explanation for the prevalence of subject relatives: the subject relative belongs to an entrenched communicative routine, the pseudo-relative or presentational-relative construction (McCawley 1981, Lambrecht 1988, 2004).

In the presentational-relative construction, a relative clause conveys an assertion that would otherwise be conveyed by a main clause, e.g., 'A friend lives next to the Cowboy's training camp' in (1). Consequently, a presentational-relative clause, unlike a restrictive relative clause, is obligatory; sentence (2), for example, does not assert 'There are a lot of people'. Based on the assumption that the prevalence of subject relatives stems from speaker's heavy reliance on the presentational-relative construction, we predict that subject-relative modifiers of objects (i.e., OS tokens), as exemplified by (1-2), will bear more linguistic hallmarks of presentational function than object-relative modifiers of objects, i.e., OO tokens like (3). In analyses of the Switchboard corpus, we confirm this expectation. First, the matrix verb in an OS token is significantly more likely to be a light verb (like *have* or *be*) than that in an OO token. Second, the nominal head of an OS token, e.g., *a friend* in (1), is significantly more likely to be indefinite than an OO nominal head. Third, when compared to OO tokens, OS tokens are significantly more likely to allow single-clause paraphrases, e.g., 'A lot of people fall into that category' in (2).

Our findings, like those of Diessel & Tomasello (2000, 2005), reveal a close connection between subject relatives and presentational function: children's first relative-clause productions contain copular matrix verbs and intransitive relative clauses. The results of the current study show that this bias is preserved in adult

speech, suggesting that speakers learn linguistic routines rather than general principles (Bybee 2001, Goldberg 2006).