Adjacency and Locality: A Constraint-Based Analysis of Complementizer-Adjacent Extraction
Ash Asudeh, Carleton University

This paper provides a new explanation of phenomena related to extraction following an overt complementizer ("that-t effects"). This normally leads to ungrammaticality (Perlmutter 1968) — see (1) — and has received numerous theoretical treatments (Bresnan 1972, Chomsky and Lasnik 1977, and many others). We adopt the theory-neutral descriptive term 'complementizer-adjacent nominal extraction' (CANE). The novel analysis stems from the Correspondence Architecture of Lexical-Functional Grammar (LFG), making formally explicit certain implicit, native relations of the architecture. The key insight is that CANE effects concern linear string adjacency, where the string is understood as part of the syntax–phonology interface.

LFG's Correspondence Architecture (Kaplan 1987, 1989) divides the form-meaning mapping into a series of simultaneously-present, discrete modules, each of which represents distinct linguistic information. The part of the architecture that is relevant here is shown in (2). C-structure represents word order, dominance and constituency, as modelled by a standard (non-tangled) tree — i.e., a phrase-structural parse of the phonological string. F-structure models more abstract aspects of syntax, such as predication and grammatical functions, null pronominals, local and unbounded dependencies, etc. F-structure is modelled as a feature structure. The ϕ correspondence function maps elements of c-structure to elements of f-structure. The theory of unbounded dependencies assumed here (Kaplan and Zaenen 1989) does not posit any null element (trace or copy) in the extraction site.

Two components of the architecture that have received little attention are at the heart of this analysis: the syntactically unparsed string and the π correspondence function from the string to the parsed c-structure. The key observation is that the relevant grammatical notion for CANE is linear adjacency, rather than structural superiority. Evidence for this is the so-called 'adverb effect' (Culicover 1993), in which an interpolated adverbial ameliorates a CANE violation; see (3). This is unexpected, because the structural relation between the complementizer and the extraction site is not affected by the interpolation. But, the adverbial does linearly separate the complementizer and the extraction site.

The native ordering relation for strings is linear precedence. This can be represented as a function on string elements, which are characterized as words; see (4). The string is therefore assumed to be 'phonologically parsed' into units. The function N (for 'next') can be used, along with the correspondence functions π and ϕ, to identify the next string element's f-structure, notated as $\succ$. The formal definition of $\succ$ (not given here for space reasons) uses only standard LFG-theoretic notions: functions and variables. Correspondence functions in the LFG architecture also have inverses. The ϕ correspondence is a many-to-one function and its inverse, ϕ^{-1}, is a relation from an f-structure to the set of c-structure nodes that map to the f-structure. F-structures may have no c-structural correspondent, e.g. null pronominals, in which case ϕ^{-1} yields the empty set. This can be used to define a notion of phonological realization for f-structures; see (5).

The CANE constraint is defined in (6). The constraint states that if the SUBJECT of the next string-element's f-structure is phonologically realized, it cannot also be an unbounded dependency function (UDF) — i.e. the SUBJECT must be locally realized. The constraint is part of the lexical entries for complementizers, as shown for *that* in (7). Cross-linguistic, dialectal and idiolectal variation (Sobin 2002) are explained as lexical variation: if the constraint is absent, CANE violations are grammatical.

The analysis makes a number of predictions. The adverb effect follows, because in this case the string element following the complementizer is the adverbial, which either has no SUBJ (3a), satisfying the constraint vacuously, or has a locally realized SUBJ (3b). A pernicious problem for structure-based accounts of CANE is what we call the Relative Clause Paradox (RCP): CANE is grammatical in relative clauses (Chomsky and Lasnik 1977, Sobin 2002); see (8). The present analysis resolves the RCP without postulating a lexical ambiguity for the complementizer *that* or postulating operations to expand (e.g., Browning 1996, Rizzi 1997) or contract (Sobin 1987, 2002) the CP layer. A *that*-relative contains a null UDF (the null relative pronoun). Therefore, the CANE constraint does not apply. In contrast, a *that*-complement containing an overt subject extraction is predicted to be ungrammatical; see (9). Also, a relative clause that contains an embedded CANE configuration is correctly predicted to be ungrammatical; see (10). An embedded VP-topic is predicted to be grammatical (11a), so long as the subject is not extracted (11b). Examples (9) and (11b) may be independently ungrammatical (e.g., in English), but the analysis also treats them as CANE violations, which makes specific predictions. For example, the account predicts that if a language has the CANE constraint but normally has no 'Doubly Filled Comp' effects, it will have an apparent DFC effect in CANE configurations.

(1)   a.   Who do you think sneezed?

        b.   *Who do you think that sneezed?

(2)   **FORM**   **MEANING**

$$\bullet \overset{\pi}{\relbar\joinrel\longrightarrow} \bullet \overset{\phi}{\relbar\joinrel\longrightarrow} \bullet \relbar\joinrel\cdots\longrightarrow \bullet$$

        string   c-structure   f-structure

(3)   a.   Who did you say that, just a minute ago, sneezed?

        b.   Who does Joan think that, with Mary out of the picture, might receive the nomination?

(4)   N: W → W, where W is the set of words in the string (string elements)

(5)   REALIZED(f) iff $\phi^{-1}(f) \neq \emptyset$, where f is an f-structure

(6)   ¬[REALIZED($\succ$ SUBJ) ∧ (UDF($\succ$ SUBJ))]

        where UDF is an unbounded dependency function (FOCUS or TOPIC)

(7)   *that*, C   ¬[REALIZED($\succ$ SUBJ) ∧ (UDF($\succ$ SUBJ))]

(8)   This is the person that sneezed.

(9)   *I know which man that saw Robin.

(10)   *This is the person who Mary thinks that sneezed.

(11)   a.   Robin knows that doubt Mary John never could.

        b.   *Who does Robin know that doubt Mary never could?

References

Bresnan, Joan. 1972. Theory of Complementation in English Syntax. Ph.D. thesis, MIT, Cambridge, MA. Reprinted as Bresnan (1979).

—. 1979. *Theory of Complementation in English Syntax*. New York: Garland.

Browning, Marguerite A. 1996. CP Recursion and *that-t* Effects. *Linguistic Inquiry* 27(2): 237–255.

Chomsky, Noam, and Howard Lasnik. 1977. Filters and Control. *Linguistic Inquiry* 8: 425–504.

Culicover, Peter W. 1993. Evidence Against ECP Accounts of the *that-t* Effect. *Linguistic Inquiry* 24: 557–561.

Dalrymple, Mary, Ronald M. Kaplan, John T. Maxwell III, and Annie Zaenen, eds. 1995. *Formal Issues in Lexical-Functional Grammar*. Stanford, CA: CSLI Publications.

Kaplan, Ronald M. 1987. Three Seductions of Computational Psycholinguistics. In Peter Whitelock, Mary McGee Wood, Harold L. Somers, Rod Johnson, and Paul Bennett, eds., *Linguistic Theory and Computer Applications*, 149–181. London: Academic Press. Reprinted in Dalrymple et al. (1995: 339–367).

—. 1989. The Formal Architecture of Lexical-Functional Grammar. In Chu-Ren Huang and Keh-Jiann Chen, eds., *Proceedings of ROCLING II*, 3–18. Reprinted in Dalrymple et al. (1995: 7–27).

Kaplan, Ronald M., and Annie Zaenen. 1989. Long-Distance Dependencies, Constituent Structure, and Functional Uncertainty. In Mark Baltin and Anthony Kroch, eds., *Alternative Conceptions of Phrase Structure*, 17–42. Chicago, IL: University of Chicago Press. Reprinted in Dalrymple et al. (1995: 137–165).

Perlmutter, David M. 1968. Deep and Surface Structure Constraints in Syntax. Ph.D. thesis, MIT, Cambridge, MA. Reprinted as Perlmutter (1971).

—. 1971. *Deep and Surface Structure Constraints in Syntax*. New York: Holt, Rinehart and Winston.

Rizzi, Luigi. 1997. The Fine Structure of the Left Periphery. In Liliane Haegeman, ed., *Elements of Grammar*, 281–337. Dordrecht: Kluwer.

Sobin, Nicholas. 1987. The Variable Status of Comp-Trace Phenomena. *Natural Language and Linguistic Theory* 5: 33–60.

—. 2002. The Comp-Trace Effect, the Adverb Effect and Minimal CP. *Journal of Linguistics* 38: 527–560.