Computation and the Justification of Grammars

Pauli Brattico

This methodological note revisits the original criteria proposed by Chomsky (1965) for the justification of grammars and suggests that modern computational methods could provide a useful tool for such purposes. Fully rigorous methods can help assessing observational, descriptive, explanatory and psycholinguistic adequacy of formally rigorous linguistic theories. The methodology is applied to the study of Finnish agreement.

Keywords: *agreement, computational linguistics, Finnish, linguistic theory, justification*

1 Introduction

It is common knowledge that the less advanced sciences, such as linguistics or sociology, do not generate the kind of cumulative knowledge characteristic of the more advanced disciplines. It is possible to accomplish productive and distinguished careers in the "soft sciences" while maintaining opposing views on virtually every issue, no matter how fundamental. In linguistics, for example, no agreement exits on questions such as what language is.[1]

One possible explanation for this discrepancy is that the more advanced sciences use a different method for justification than the less advanced ones. In the former, theories are connected to observations by deductive calculation. This system, first used in its present form by Galileo and later institutionalized by Newton, possesses an unrivaled epistemological power because it removes opinion from scientific justification. The medieval criteria of thought experiment, common sense, human intuition, authority, popularity, institutional structure, author reputation, political correctness, sociology of science, imagination, or any type of Augustine's "divine illumination" play no role in justification in these fields (although they do play a role in other affairs such as discovery). Thus, in linguistics, too, we should aim to "construct a formalized general theory of linguistic structure" because by "pushing a precise but inadequate formulation to an unacceptable conclusion, we can often expose the exact source of this inadequacy and, consequently, gain a deeper understanding of the linguistic data. More positively, a formalized theory may automatically provide solutions for many problems other than those for which it was explicitly designed" (Chomsky 1957: 5). In fact, a grammar that is "perfectly explicit" and does not rely "on the intelligence of the understanding reader" (Chomsky 1965: 4) is considered to be "generative," hence the term "generative grammar"

1 It is possible to accomplish distinguished professional careers in linguistics by believing that language is not a natural phenomenon at all or that it is a biological property of the human brain; that it is based on innate properties of the human cognitive architecture or that it is learned by simplest Pavlovian association; that it has autonomous syntax or that its principles are covered by general cognitive or even pragmatic principles; that it has recursive syntax or only nonrecursive components; that is has no qualitative differences when compared to nonhuman 'languages' or that the human language is biologically unique; that there is a specialized language faculty or none exists, and so on. Virtually any imaginable position can be and has been entertained despite the fact that the data remains the same.

refers to a theory that is formal in this exact sense. Yet, such methods are almost never employed, perhaps due to the complexity of the required linguistic calculations making them unfeasible from the point of view of practical research projects. I propose in this note that modern computational tools provide a feasible way out of this methodological difficulty.

2 Background

First we must define the term "computational linguistic theory" to dispel some myths. A computational linguistic theory must satisfy two conditions: it must be (i) unambiguous and (ii) expressed in a machine-readable way. The requirement that a scientific theory must be unambiguous means that it does not rely on notions or assumptions that are open to interpretation. This allows the researcher to connect the theory with observation in a way that does not leave room for opinion, disagreement or logical gaps. In addition, when scientists put forward ambiguous theories they must be implicitly or explicitly assuming that an unambiguous formulation exists; believing to the contrary would be tantamount to saying that the theory must have some poetic quality making it necessarily ambiguous. Condition (i) is therefore nonnegotiable. Condition (ii) imposes an additional requirement: the theory must be provided in a machine-readable format or at the very least implemented in such notation. This allows the researcher to use a computer to test the theory against observation by using deductive calculation. In short, a computational linguistic theory is an ordinary linguistic theory formulated in some unambiguous notation that a machine can understand. No other properties are at stake.

While I will claim that linguistics can benefit from the use of rigorous computational methods in justification, this is also the only thing I want to claim. I do not propose to eliminate human intuition from the scientific discovery process or from any subject matter consideration, or to replace the 17th century scientific method with 19th century positivism that suspended all abstract theorizing. My concern is justification: how to bridge the theory, discovered by whatever mystical process, with observations, acquired by some means I do not want to restrict. Similarly, the point is not to replace linguistic theorizing with data mining or apply computational techniques to datasets in the hopes of discovering surface correlations. While computational discovery procedures can be useful in some contexts, they are irrelevant to the matters discussed in this article. Finally, the medieval method that relies on divine illumination or some other form of superior human cognitive capacity in justifying scientific hypotheses is, whatever faults it has, able to generate true theories. It is also able to produce interesting observations. One can discover groundbreaking truths even by pure luck. What the medieval method is unable to produce is agreement.

To show that the proposed computational methodology is feasible within the context of a real linguistic research project, I will use a concrete computational framework in this article as an example. The example system is a Python-based program that I wrote to provide an idealized "brain model" of a speaker of any language allowing the researcher to embed it with linguistic analyses and to test them by calculation. The framework consists of several interconnected components, the most important shown in Figure 1.

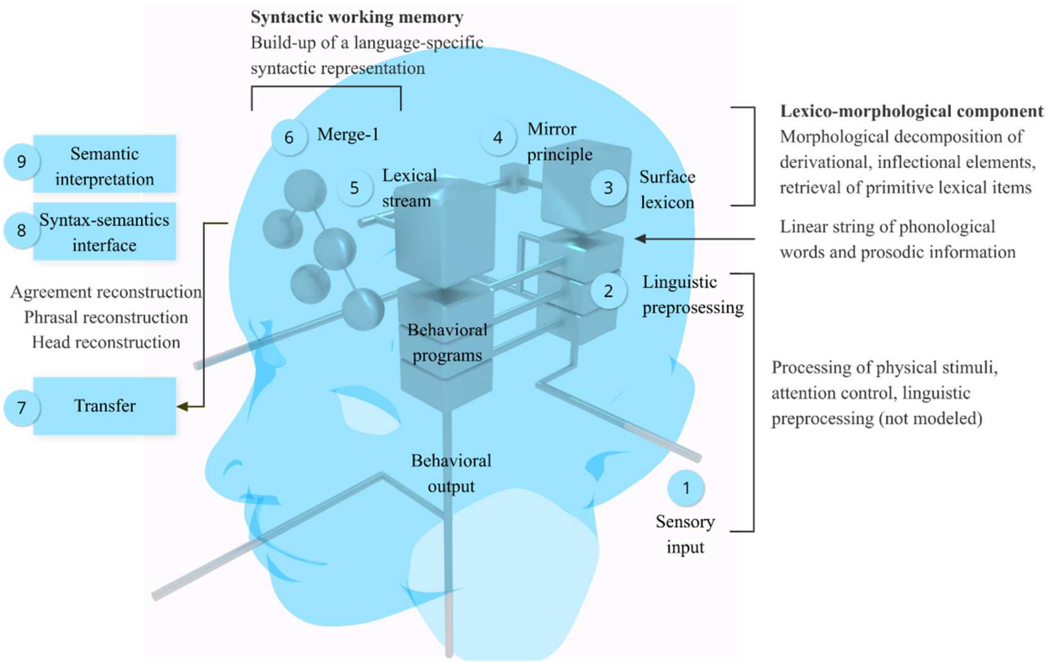

Figure 1: *Diagram of the various components of the Python-based brain model used as an example in this article. Input sentences are linear sequences of phonological words (1–2), which are processed by a lexico-morphological component performing lexical retrieval and morphological decomposition (3), followed by mapping of lexical items into syntax (4–5), generation of parsing solutions (6) and transfer (7–8) into systems of semantic interpretation (9)*

The model reads input sentences from left to right, retrieves each input word from the lexicon (3, Figure 1), merges them into a partial phrase structure in the current active working memory (6), and, once all words have been consumed from the input, transfers the calculated result to the syntax-semantics interface (7, 8) for evaluation and semantic interpretation (9).[2] It therefore maps phonological input sentences from the sensory interface(s) into sets of syntactic analyses and semantic interpretations. Input sentences that are judged ungrammatical are not interpreted semantically and are marked ungrammatical in the calculated output. The architecture was developed on the basis of earlier work by Phillips (1996) and is documented in Brattico (2019a).

Consider an input sentence *the horse raced past the barn*. The sentence is consumed one word at a time from left to right (1a) while each lexical item, retrieved on the basis of the phonological word in the input, is merged incrementally to a partial syntactic representation that exists in the algorithm's working memory (1b–c).

(1)	a.	*the*	*	*horse*	*	*raced*	*	*past*	*	*the*	*	*barn*	(Input)
	b.	↓		↓		↓		↓		↓		↓	(Merge)
	c.	[[the		horse]		[raced[		past		[the		barn]]]]	(Result)

[2] It is usually assumed in the cognitive sciences that the human conceptual system is not language-specific. This presupposes that the syntactic processing pathway eliminates language-specific features from the input before interpretation. Transfer (7, Figure 1) accomplishes this task. It incorporates a reverse-engineered chain creation algorithm, thus much of the linguistic theorizing currently in focus in generative theorizing is encapsulated inside this component.

If no more words appear, (1c) is interpreted semantically. Suppose however that there is one more word, an intransitive verb *fell*, in the input. Because there is no grammatically legitimate position for the intransitive verb in (1c), the algorithm reconsiders earlier parsing decisions by backtracking. Ambiguities are discovered in the same way. Backtracking finds all possible grammatical analyses and semantic interpretations for any given input string that are consistent with the linguistic hypotheses incorporated into the model. Ungrammatical input sentences are marked in the output as such. We examine this process in Section 4.

Whether this particular model is correct or even plausible is not important. Most linguists would judge its underlying assumptions as misguided. An expert who reviewed a manuscript advocating the above model criticized it as completely clueless, thus "nothing whatsoever" justified it, in his or her opinion. There was no "conceptual realm […] in which it might make some sense or have some application", and furthermore there was no explanation, according to this expert, of "who or what is supposed to carry out these operations". The idea that language is involved with neurocognitive computations of some sort was considered so outrageous as to be incomprehensible. In that reviewer's expert opinion, then, in order to proceed "we need to be looking at areas of linguistic inquiry […] very far removed from anything this author is interested in". This complete lack of agreement on every aspect of every theory aside (a standard feature of the less advanced sciences), my point is not to argue whether the model is correct or incorrect; rather, the point is that it is not justified by the type of subjective opinion exhibited by the reviewer in these remarks. Let us examine how it is justified.

3 Justification of grammars

The model described in Section 2 maps input sentences into sets of semantic interpretations and, if no mapping is found, judges them ungrammatical. It therefore captures a notion of linguistic competence by partitioning any set of input sentences into grammatical and ungrammatical, and by providing the former with a grammatical analysis (or several analyses). The Python implementation makes this process automatic.

Let us consider the word order study reported by Brattico (2020). Twenty Finnish seed sentences were selected that represent basic construction types in Finnish. All possible word order permutations were mechanically generated from the seed sentences. This method generated 119800 unique word orders. We can regard this set as a minimal word order corpus for this language, containing all possible word order permutations derived from a set of relatively simple seed constructions. Working with a corpus of this size with the traditional paper-and-pencil methodology would be infeasible, but it presents a trivial task for the computer. The recognition algorithm enriched with grammatical word order principles used less than a day to calculate structural analyses and grammaticality judgements for each sentence in this corpus. The model was justified, to the extent that it was, by matching the calculated output with native speaker judgments.

Let us break the process into several stages. I build an ad hoc test corpus of 2038 construction types that represent a wide variety of linguistic constructions in Finnish, English and Italian. The test corpus, therefore, contains data considered relevant to the subject matter under study. If we were concerned with word order, then all logically possible word order permutations from some set of basic constructions should appear in this file. If the focus was on some specific phenomenon such as *pro*-drop, we should use finite clauses exemplifying possible and impossible *pro*-drop configurations. If the study aims for establishing crosslinguistic generalizations, then we include sentences from several languages. The contents of the test corpus used for the purposes of the present article are listed in Table 1.

GROUP	DESCRIPTION
1	**Basic construction types**
1.1	Basic verb classes
1.1.1	Intransitive verbs
1.1.2	Transitive verbs
1.1.3	Ditransitive verbs
1.1.4	Ditransitive verbs plus PP argument
1.1.5	Two PP arguments
1.2	Special finite elements
1.2.1	Auxiliary-like negation
1.2.2	Modal constructions, Neg + Modal + V
1.2.3	Pure tensed auxiliaries
1.3	Clausal infinitivals
1.3.1	Clausal infinitivals in English
1.3.2	English OC-constructions
1.3.3	Finnish clausal infinitivals (A/inf, VA/inf)
1.4	Nominals
1.4.1	Basic DP constructions
1.4.2	N + clausal infinitival
1.5	Adpositions
1.5.1	Prepositions, postpositions
1.6	Embedded *that*-clauses
1.6.1	Embedded *that*-clauses in Finnish and English
1.7	Relative clauses
1.7.1	Subject relativization
1.7.2.	Object relativization
1.8	Lexical ambiguity
1.8.1	Lexical ambiguity tests (frequency based)
2	**Adjuncts and adjunction structures**
2.1	PP adjuncts
2.1.1	Postverbal PP adjunct constructions
2.2	Adjectives
2.2.1	DP-internal adjectives (Finnish, English)
2.3	Clausal adjunct infinitivals in Finnish
2.3.1	MA-infinitivals
2.3.2	ESSA-infinitival
2.3.3	TUA-infinitival
2.3.4	E-infinitival
3	**A-bar (operator) movement and pied-piping**
3.1	Basic interrogatives
3.1.1	Subject and object interrogatives
3.2	Pied-piping
3.2.1	Pied-piping in Finnish and English
3.3	Islands
3.3.1	*CED-effects
3.3.2	*Extraction from embedded wh-clause
3.3.3	*Extraction from DP
3.3.4	*Extraction from embedded subject position
3.4	Left-peripheral C-features (Finnish only)
3.4.1	All agglutinative combinations
3.4.2	Single C-feature (subjects, objects)

3.4.3	*Double filled operator position
3.4.4	C-features in connection with pied-piping
3.4.5	C-features and noncanonical word order
3.4.6	C-features, pied-piping and noncanonical order
3.5	Embedded interrogative clauses
3.5.1	Canonical embedded interrogatives
3.5.2	Noncanonical embedded interrogatives
3.5.3	Selection tests (main verb + embedded clause)
3.6	Operator in situ (wh and focus)
3.6.1	Wh in situ (echo interpretation)
3.6.2	Prosodic focus in situ
3.6.3	In situ in embedded clause and pied-piping
3.6.4	*Ungrammatical in situ constructions
4	**Case assignment**
4.1	Finite clause, nominative and partitive
4.1.1	Canonical clause, nominative and partitive
4.2	Finite clause, accusative
4.2.1	Canonical accusative configuration
4.2.2	Accusative in the scope of negation
4.2.3	Accusative and agreement
4.2.4	Long-distance accusative effects
4.3	Adpositions and case
4.3.1	Adpositions and postpositions
4.4	Infinitivals and case
4.4.1	Genitives and partitive objects
4.5	Possessive construction
4.5.1	D + poss(DP) + N
4.6	Numeral construction
4.6.1	Two numeral types
4.7	Adverbials, direct object marking
4.7.1	MALLA-adverbial
4.8	Case marking on DP-adverbials
4.8.1	Accusative and partitive alteration
4.9	Special constructions
4.9.1	Psych-verb construction
4.9.2	Impersonal passive
4.9.3	Copula
5	**Agreement**
5.1.1	Standard finite S-V agreement
5.2.1	Standard S-V agreement with noncanonical order
5.2.2	Incorrect agreement with noncanonical order
6	***Pro*-drop (null subject)**
6.1	Finite *pro*-drop
6.1.1	Finite *pro*-drop in Finnish and Italian
6.2	Finite *pro*-drop with noncanonical order
6.2.1	*Pro*-drop with noncanonical order (Finnish)
6.3	Third person *pro*-drop in Finnish (partial drop)
6.3.1	With and without long distance antecedent
7	**Control**
7.1	Partial control in Finnish
7.2	Standard control
7.2.1	Want-class
7.2.2	OC-construction
7.2.3	Anti-OC construction

9.6.1	Intransitives
9.6.2	Transitives
9.6.3	Ditransitives
9.6.4	Neg-to-C
9.6.5	LHM
10	**Clitics (Italian)**
10.1	Direct object clitics
10.2	Two-verb constructions
10.3	Three-verb constructions
10.4	Clitic agreement constructions and tests
10.5	Indirect clitic arguments
10.6	Clitic clusters
10.7	Restructuring
10.8	Reflexives

Table 1: *Test sentences (2038 in total). A category that exemplifies only ungrammatical sentences is marked with an asterisk*

One strength of this framework is that all hypotheses and theories aspiring to explain some linguistic phenomena can potentially agree to a common dataset, as defined by the test corpus. Another benefit is that everybody will come to the arena with the same requirement: propose a formula that calculates the same data. We will also eliminate a situation where two linguistic theories compete against each other while working, implicitly or explicitly, with different datasets.

Next, a script was deployed that read all sentences from the test corpus and fed them to the idealized speaker model (Figure 1), which then processed the sentences on the basis of the linguistic principles hypothesized by the author. In this way, we can examine if the hypothesis replicates the grammaticality judgments of human informants and "presents the observed data correctly" (Chomsky 1964: 28). This constitutes a minimal criterion for any scientific hypothesis, in any field. To do this, we create a gold standard and compare it with the calculated output. An example comparison, when I ran the test corpus through an algorithm that existed at the time of this writing, is provided in Figure 2. The gold standard is on the left, model output is on the right. The rightmost yellow column shows the comparison over the whole test corpus. Discrepancies are highlighted in red.

If a sentence is judged ungrammatical, then no output apart from the judgment itself is produced. Ungrammatical sentences have neither well-defined phrase structure representations nor semantic interpretations. To find out why some sentence was judged ungrammatical, we consult a derivational log file that stores all linguistically relevant computational steps executed during the calculations. Let us examine the expression *se talo* 'that.NOM house.NOM' (#197, Figure 2, line 312) that the model judged wrongly as ungrammatical. We locate the input from the derivational log file and examine what happened when the model processed that input. This is shown in Figure 3.

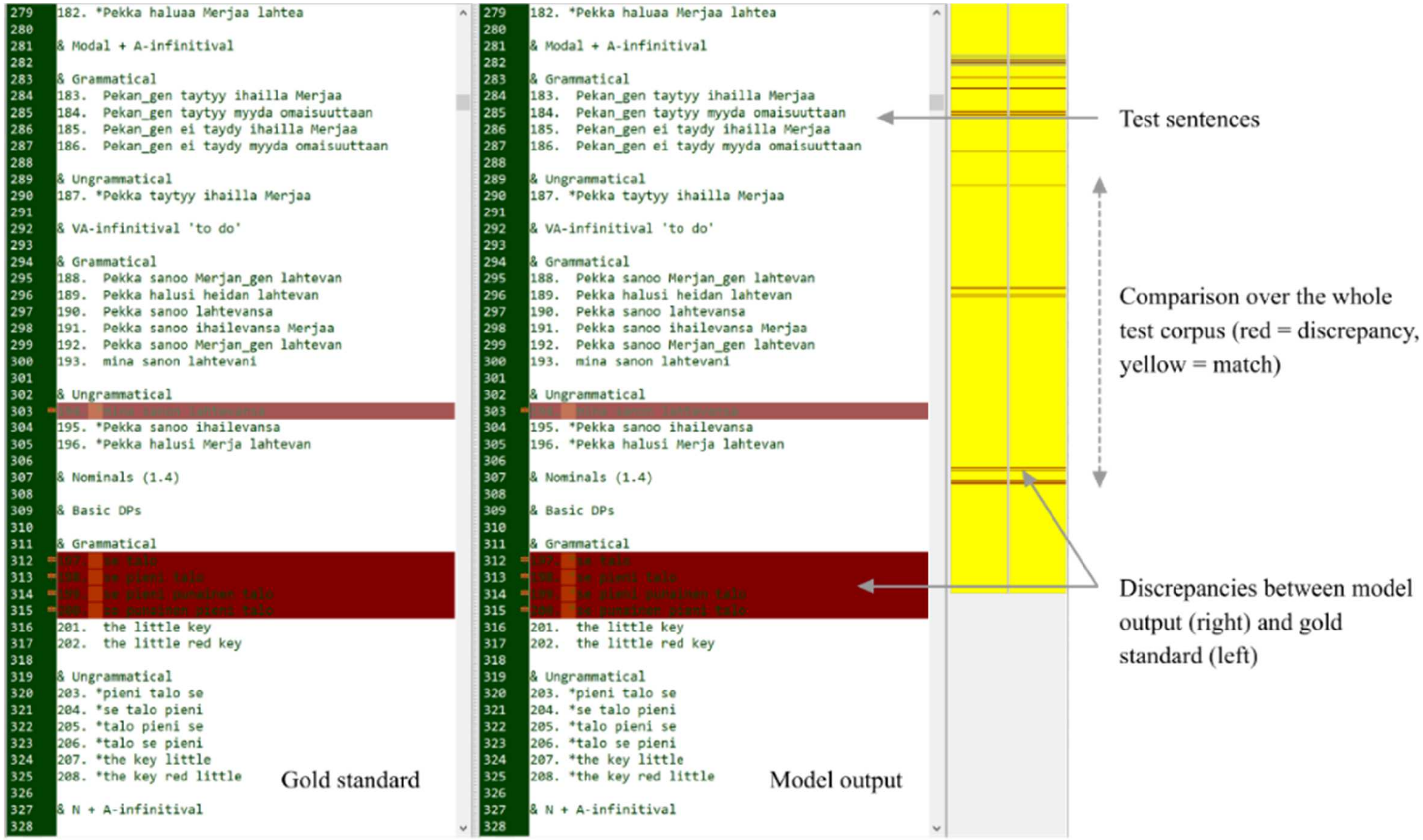

Figure 2: *A comparison between the gold standard (native speaker output, left) and calculated modal output (right). Discrepancies are marked by red lines by the automatic comparison tool. The rightmost yellow column shows the comparison over the whole test corpus*

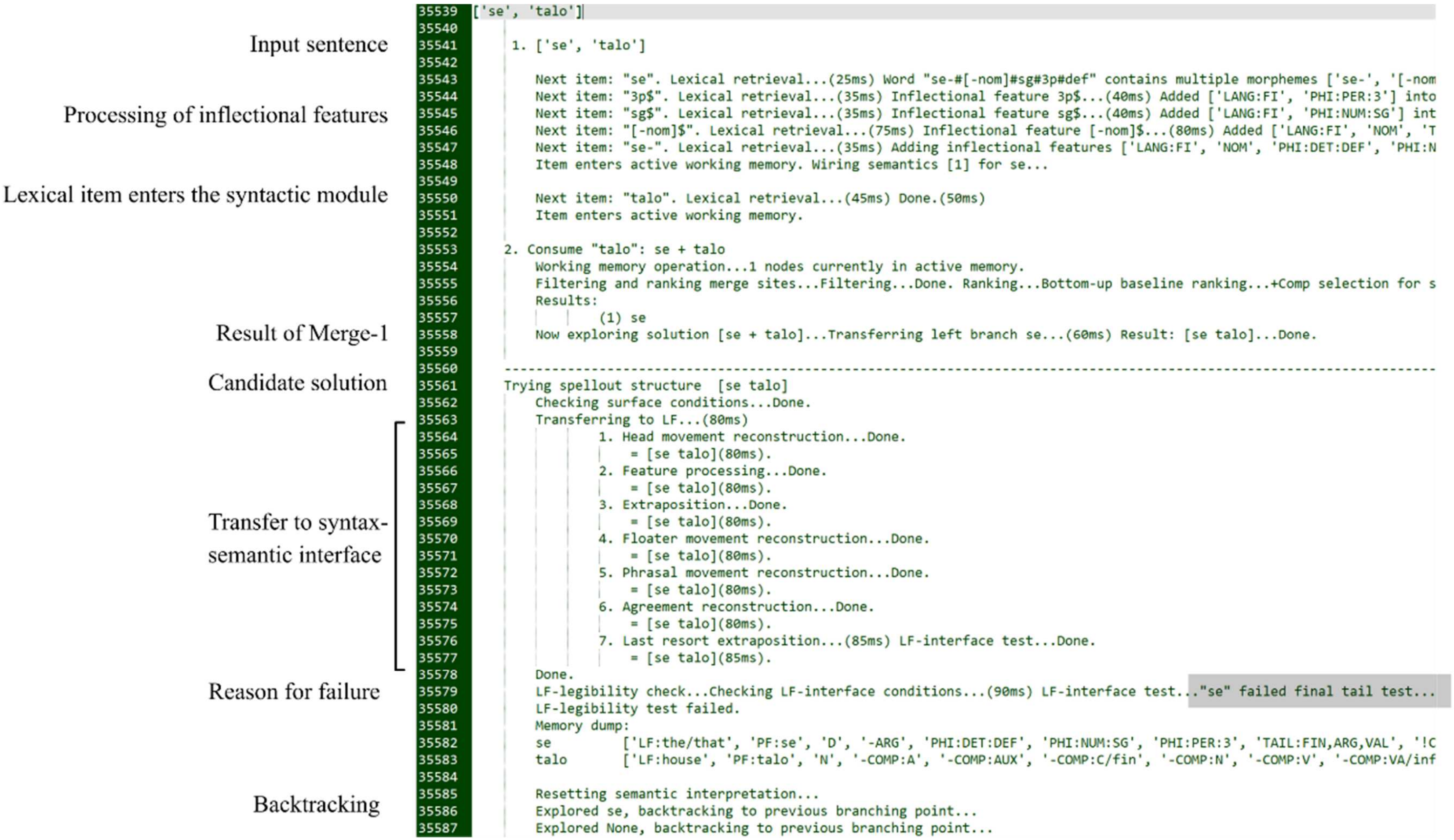

Figure 3: *Screenshot from the derivational log file showing the derivation of an isolated DP* se talo *'that.*NOM *house.*NOM*'*

The file is read in top-down order. The first element *se* 'the.NOM' is processed on Step 1 (lines 35541–35551), followed by the processing of *talo* 'house' (Step 2, lines 35553–35558). They are merged together to form [$_{DP}$ *se talo*] (line 35558), which is transferred (lines 35563–35578) to the syntax-semantic interface. The derivation fails at the syntax-semantics interface because the nominative case feature of *se* 'that.NOM' could not be checked: the required clausal context was missing (line 35579). The model tried to backtrack (lines 35586).

The hypothesis is revised until the model and data match. Once they do, the model is said to be *observationally adequate*. No natural language syllogisms or intuitive jumps occur in the justification, and no authority is allowed to use Augustine's divine illumination to consider that "nothing whatsoever is said by way of justifying this analysis" or that it does not make "any sense" or does not have "any application". In fact, any defect, problem or limitation is completely transparent. They are shown in Figure 2.

Suppose we have a hypothesis that is observationally adequate or nearly observationally adequate. This requirement alone is not sufficient, or even particularly interesting. All it says is that the algorithm captures something about the formal "shape" of the data. One can reach observational adequacy by storing the whole test corpus into the algorithm's memory. A trivial model of this kind does not contribute anything to any linguistic theory, in the same was as a table-lookup catalog of precomputed values or measurements used in engineering does not contribute anything to physics. Explanations which are relevant within the context of a linguistic theory are said to have *descriptive adequacy* (Chomsky 1965).

What counts as a descriptively adequate grammar depends on the linguistic framework. A linguist who has constructed an observationally adequate theory by using a connectionist model will judge the matter differently than another researcher who works with modern minimalism. Similarly, a linguist who begins from a semantic-based dependency grammar will have different concerns than one who adopts some variation of cognitive linguistics. The expert authority who judged the present model as grossly incorrect proposed that we should use a historical method to capture the same data. In general, then, descriptive adequacy is meaningful in relation to "grammars that are paired with some linguistic theory" (Chomsky 1964: 52). This is because it is "always possible to describe the linguistic intuition of the native speaker in a completely ad hoc way in any particular case if we drop the requirement that the grammar be constructed in accordance with some fixed model or if we allow the associated linguistic theory to be completely general and without content" (ibid.).

To assess descriptive adequacy, we examine the calculated output in the context of an existing linguistic theory. Let us consider the output in connection with an interrogative clause *ketä Pekka ihaile-e* 'who.PAR Pekka.NOM admire-PRS.3SG' (sentence #383 in the test corpus). The calculated output is shown in Figure 4. The input sentence is associated with a syntactic interpretation (line 9435) and illustrated further in Figure 5 generated by the algorithm.

Input sentence
Calculated output
Semantic interpretation
Performance metrics
Discourse inventory (language-external objects)

```
383. keta Pekka ihailee

    [[D kuka]:1 [C [<D Pekka>:2 [T [<__>:2 [v [__:1 ihaile-]]]]]]]

    Semantics:
    Recovery: ['Agent of ihaile-(Pekka)', 'Agent of v(Pekka)']
    Aspect: []
    D-features: [('[D kuka]', 'D:WH', '1')]
    Operator bindings: [('[D kuka]', 'C')]
    Speaker attitude: ['Interrogative']
    Information structure: {'Marked topics': ['2'], 'Neutral gradient': ['3'], 'Marked focus': []}

    Resources:
    Total Time:1070, Garden Paths:0, Memory Reactivation:0, Steps:9, Merge:6, Move Head:8, Move Phrase:2,
    A-Move Phrase:0, A-bar Move Phrase:1, Move Adjunct:1, Agree:1, Phi:3, Transfer:8, Item streamed into syntax:7,
    Feature Processing:0, Extraposition:2, Inflection:8, Failed Transfer:4, LF recovery:2,
    LF test:13, Filter solution:4, Rank solution:3, Lexical retrieval:15, Morphological decomposition:3,
    Mean time per word:356, Asymmetric Merge:26, Sink:6, External Tail Test:18,

    Discourse inventory:
    Object 1 ['§Thing']
        Referring constituent: D
        Order gradient: 1
        Reference: [D kuka]
        Semantic type: ['§Thing']
        Operator: True
        Bound by: C
        Operator interpretation: ['Interrogative']
    Object 2 ['§Thing']
        Referring constituent: D
        Order gradient: 2
        Reference: [D Pekka]
        Semantic type: ['§Thing']
        Operator: False
        Marked gradient: High
        In information structure: True
    Object 3 ['§Proposition']
        Referring constituent: C
        Order gradient: 3
        Reference: [[D kuka] [C [<D Pekka> [T [<D Pekka> [v [[D kuka] ihaile-]]]]]]]
        Semantic type: ['§Proposition']
        Operator: False
        In information structure: True
```

Figure 4: *Calculated output for* Ketä Pekka ihaile-e? *'who.PAR Pekka.NOM admire-PRS.3SG'*

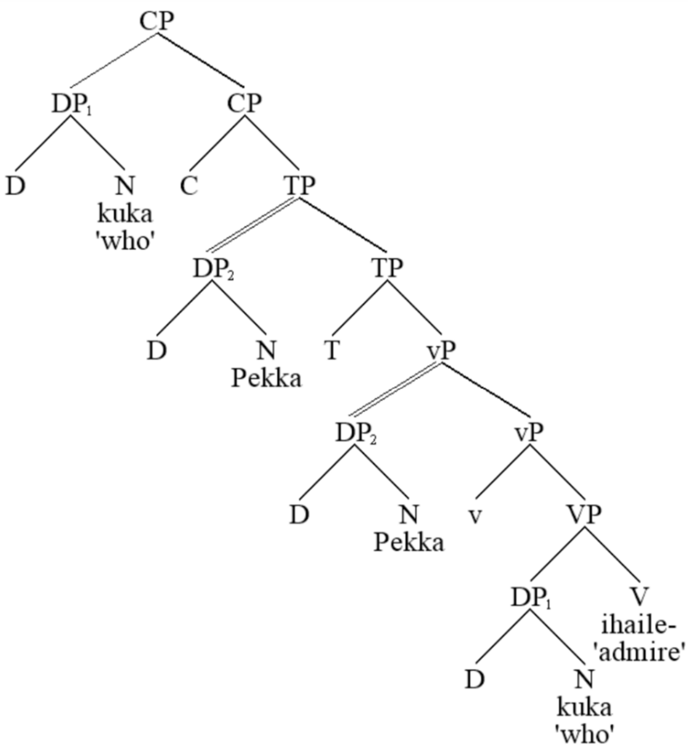

Figure 5: *The phrase structure analysis calculated by the model for the input sentence #383*

The researcher must determine whether the calculated syntactic interpretation is correct and/or plausible, given the background theory and the independent evidence that has motivated it. In this case it is: the interrogative pronoun was reconstructed correctly to the object position, while the grammatical subject/topic was reconstructed to the thematic SpecvP position. From the point of view of the theoretical framework used in this particular theory this is a plausible output.

The fact that descriptive adequacy must be relativized to a background theory or framework does not mean that no progress is possible beyond this point or that anybody can believe anything given any background theory. Let's consider, as an example, the fact that the algorithm analyzed the sentence as consisting of hierarchically organized parts. Suppose we change the algorithm so that it replaces hierarchical structures with flat representations. We will also write a function that interprets flat representations. What motivates any of these analyses if observational adequacy can be reached by either model?

The structure-dependence principle is based on linguistic evidence indicating that the object and the verb constitute one unit that does not include the subject. We could therefore claim that it is supported by empirical evidence. The problem, though, is that this justification is not visible in the calculations. Structure dependence can be justified for example by relying on reflexive binding, yet reflexive binding was not part of the algorithm and there were no reflexives in the test set. Therefore, the force one is willing to grant to this argument is to a degree subjective and, consequently, a significant portion of professional linguists remain skeptical towards these claims. One is free to assert by divine illumination or by some other superior cognitive ability that the same properties could be explained by a number of competing but simpler models, such as by purely historical analysis—or indeed that they could not.

Yet, we can now evade the whole issue. Suppose a researcher develops a model of Finnish word order that generates flat structures instead of hierarchical ones and provides a mechanism associating each flat structure with a semantic interpretation. If it reached observational adequacy and generated correct semantic interpretations, we would have two competing models that cannot be distinguished from each other in terms of their predictive success, one that uses hierarchical representations and another that projects flat structures. This would not be a problem, though, because eventually both researchers must face the evidence relevant for deciding the issue. We proceed by generating an agreed-upon dataset deemed relevant to the issue, thus it should contain at least reflexive data, control constructions, subject–object extraction asymmetries, and other data pertinent to this issue, and examine how each model handles that set. If the matter cannot be settled by this method, on the other hand, then the two models must be judged notational variants of only one underlying model—not an unusual situation and in no way a barrier to progress in the advanced sciences.[3]

Let us consider semantic interpretation next. Figure 6 shows part of the calculated semantic output for a simple interrogative clause.

```
9433 383. keta Pekka ihailee
9434
9435      [[D kuka]:1 [C [<D Pekka>:2 [T [<__>:2 [v [__:1 ihaile-]]]]]]
9436
9437      Semantics:
9438      Recovery: ['Agent of ihaile-(Pekka)', 'Agent of v(Pekka)']
9439      Aspect: []
9440      D-features: [('[D kuka]', 'D:WH', '1')]
9441      Operator bindings: [('[D kuka]', 'C')]
9442      Speaker attitude: ['Interrogative']
9443      Information structure: {'Marked topics': ['2'], 'Neutral gradient': ['3'], 'Marked focus': []}
```

Figure 6: *Part of the calculated semantic output for a basic interrogative sentence. There is an error in the calculation; see the main text for explanation*

Line 9438 shows that the model interprets Pekka as the agent of the whole event ('who admires') but wrongly interprets Pekka as an argument of the verbal stem 'admire'. The latter should have been 'who'. The sentence does not mean 'Pekka admires himself'. Examination of derivations of other sentences in the test corpus reveals that this problem has to do with the fact that the patient is an interrogative operator: regular direct objects are interpreted correctly. But in this case the model output does *not* match with the semantic intuition of a native speaker, and some correction is needed. The rest of the interpretation appears to be correct.

The semantic interpretation shown in Figure 6 pools various aspects of semantic interpretation calculated during processing. The researcher can populate this structure with anything deemed relevant for a particular research agenda. Ideally, we would like to base the calculations on a more principled semantic system. Some initial steps towards such an explanation are taken by using a data structure called discourse inventory, shown in Figure 7.

[3] A well-known example of this is Freeman Dyson's proof of the equivalence of Feynman's and Tomonaga-Schwinger's approaches to quantum electrodynamics (Dyson 1949).

```
Discourse inventory:
Object 1 ['§Thing']
    Referring constituent: D
    Order gradient: 1
    Reference: [D kuka]
    Semantic type: ['§Thing']
    Operator: True
    Bound by: C
    Operator interpretation: ['Interrogative']
Object 2 ['§Thing']
    Referring constituent: D
    Order gradient: 2
    Reference: [D Pekka]
    Semantic type: ['§Thing']
    Operator: False
    Marked gradient: High
    In information structure: True
Object 3 ['§Proposition']
    Referring constituent: C
    Order gradient: 3
    Reference: [[D kuka] [C [<D Pekka> [T [<D Pekka> [v [[D kuka] ihaile-]]]]]]]
    Semantic type: ['§Proposition']
    Operator: False
    In information structure: True
```

Figure 7: *A screenshot of the output containing contents of the discourse inventory created during the processing of a simple interrogative clause*

The discourse inventory is populated with language-external semantic objects during the derivation. The term "semantic object" refers to (mental representations of) language external objects, such as persons, propositions or events that are denoted by the linguistic expressions that occurred in the input clause.

In short, then, descriptive adequacy is evaluated by comparing the model output against semantic and syntactic intuitions of native speakers and/or against a theoretical matrix as defined by a larger linguistic background theory.

Explanatory adequacy refers to a further requirement which demands that the model agrees with external evidence concerning language acquisition. Thus, a grammar that satisfies the condition of explanatory adequacy provides "a principled basis, independent of any particular language, for the selection of the descriptively adequate grammar of each language" (Chomsky 1964: 29), where "selection" refers, or can refer to, learning. A model of this type describes a language acquisition device that maps sensory data available to a language learner into a descriptively adequate grammar, relying on "innate predisposition of the child to develop a certain kind of theory to deal with the evidence presented to him" (Chomsky 1965: 26).

Any language comprehension algorithm will, as a matter of necessity and independent of whether it models language learning or not, contain fixed and variable parts. How the division is implemented is an empirical question. The algorithm used here assumes that while the computational principles remain universal, lexicons may differ. A model of this kind reaches explanatory adequacy if and only if it captures observations from several (or, in an ideal sense, all) languages. Explanatory adequacy can therefore be assessed by using several languages in the test corpus. In other words, such a theory must "develop an account of linguistic universals that […] will not be falsified by the actual diversity of languages" (Chomsky 1965: 28).

Psycholinguistic adequacy refers to the condition that the model should not contradict anything known independently concerning real-time language processing. To move towards such a goal, we can give the program an ability to monitor its own performance.

Thus, the results file contains a segment providing performance metrics, reproduced in Figure 8.

```
Resources:
Total Time:1070, Garden Paths:0, Memory Reactivation:0, Steps:9, Merge:6, Move Head:8, Move Phrase:2,
A-Move Phrase:0, A-bar Move Phrase:1, Move Adjunct:1, Agree:1, Phi:3, Transfer:8, Item streamed into syntax:7,
Feature Processing:0, Extraposition:2, Inflection:8, Failed Transfer:4, LF recovery:2,
LF test:13, Filter solution:4, Rank solution:3, Lexical retrieval:15, Morphological decomposition:3,
Mean time per word:356, Asymmetric Merge:26, Sink:6, External Tail Test:18,
```

Figure 8: *Performance metrics provided by the algorithm when processing a simple intransitive clause*

Total time refers to the predicted processing time in milliseconds, thus the model predicts that it takes an average hearer approximately one second to process the sentence *ketä Pekka ihaile-e?* 'who.PAR Pekka.NOM admire-PRS.3SG'. These numbers are created by associating the linguistically relevant computational operations postulated by the algorithm with a predicted processing cost in milliseconds and then summing them over the course of the whole derivation. Predicted costs should ultimately be determined on the basis of well-established physiological properties of human neuronal information processing and then tested in laboratory experiments. They can also be used to assess the computational cost and psycholinguistic reality of any proposed linguistic model.

Performance metrics are also provided in a comma-delimited form that can be processed by external programs such as Excel, SPSS, or by Python data processing scripts as I did below. The file lists each input sentence associated with the performance metrics shown above, all written to the same line. We can group the input sentences from this file on some basis, say by using the classification present in the test corpus in Table 1. These groups can then function as independent variables; for dependent variables, we take whatever metrics are of particular interest. Figure 9 shows the results when I examined the mean predicted processing times per word as a function of the main linguistic category in the test corpus.

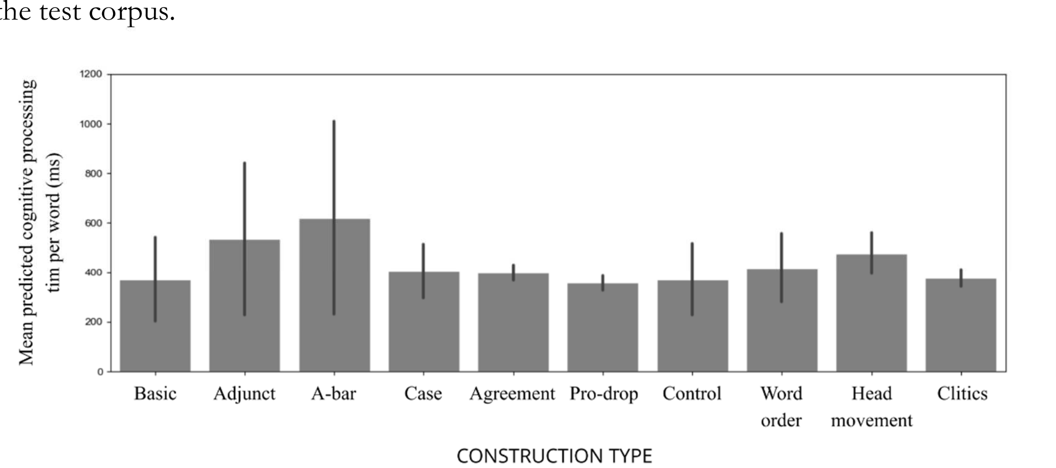

Figure 9: *Mean predicted processing times (in milliseconds) as a function of construction type. Errors bars represent standard deviation*

The first category "Basic" (Group 1 in the test corpus, Table 1) can be taken as an overall estimation of how the model performs with standard clause types. A word is predicted to take an average of 372 ms to process. The model has difficulties with sentences that involve adjunction (mean 535 ms) and operator movement (mean 619 ms). This is due to garden pathing, as shown by Figure 10.

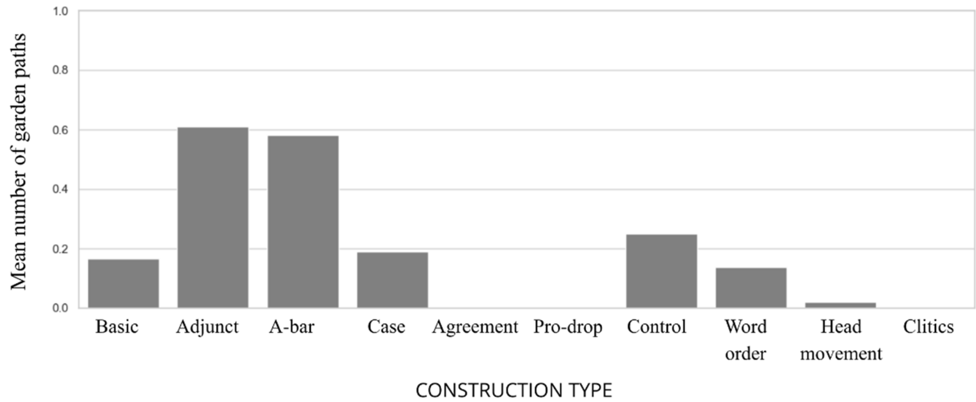

Figure 10: *Mean number of garden paths as a function of major category in the test corpus*

Whether performance properties of this kind are included into the study is for the researcher to decide.

4 Finnish agreement and computation

I will conclude this note by applying the methodology to the study of Finnish agreement to illustrate how it works in connection with a concrete empirical problem.

Finnish finite elements agree in number and person with nominative grammatical subjects (2a–e). The category of Finnish finite elements contains finite verbs (2a, e), auxiliaries (2b), negation (2c), and special modal verbs (2d).[4]

(2) a. *Jari ihaile-e naapure-i-ta-an.*
Jari.NOM admire-PRS.3SG neighbour-PL-PAR-PX/3SG
'Jari admires his neighbors.'

b. *Jari o-n ihail-lut naapure-i-ta-an.*
Jari.NOM be-PRS.3SG admire-SG.PST.PRTCPL neighbour-PL-PAR-PX/3SG
'Jari has admired his neighbors.'

c. *Jari e-i ihail-lut naapure-i-ta-an.*
Jari.NOM not-3SG admire-SG.PST.PRTCPL neighbour-PL-PAR-PX/3SG
'Jari did not admire his neighbours.'

d. *Naapure-i-den täytyy pitä-ä ovi lukossa.*
neighbour-PL-GEN must.0 keep-A/INF door in.locked
'The neighbours must keep the door locked.'

[4] Abbreviations: 0 = no agreement or default third person agreement; 1, 2, 3 = person features; A/INF = A-infinitival; GEN = genitive case; NOM = nominative case; PAST = past tense; PAR = partitive case; PL = plural; PRS = present tense; PRTCPL = participle; PST = past tense; PX = infinitival agreement marker; SG = singular.

e. *Mei-tä väsyttä-ä.*
we-PAR feel.tired-PRS.3SG
'We feel tired.'

The algorithm gets syntactic agreement configurations from the parser that it must check against possible phi-feature mismatches between the subject and finite element. This checking relation must involve finite elements and nominative thematic subjects (2a–c), and it must ignore non-nominative arguments, independent of their positions (2d–e). Checking cannot be performed presyntactically, it must access syntactic notions such as 'subject' and 'finite element'. In addition, it is not in my view possible to perform these operations with a partial phrase structure whose final structural properties remain unknown. Another option is to perform the required checking operations in some postsyntactic semantic component, but the empirical evidence (e.g., (2a–e)) indicates this to be unlikely. These considerations suggest that agreement checks are performed between the parsing stage and semantic interpretation, therefore during transfer. Accordingly, let us assume that transfer contains an agreement reconstruction operation which verifies that there are no agreement mismatches in the phrase structure produced by the parser and prior reconstruction operations.

To sharpen this idea into a formal model, we consider first a tentative agreement hypothesis that is close to the standard agreement operation assumed in virtually all current minimalist literature on Finnish (e.g., Koskinen 1998, Holmberg & Nikanne 2002, Manninen 2003, Brattico 2019b) and on UG (Chomsky 2000, 2001, 2008). According to this analysis, finite elements agree in number and person with a local DP at the thematic SpecvP position (3). The DP is reconstructed from the preverbal topic position SpecTP during transfer by a reverse-engineered phrasal movement called phrasal reconstruction.

(3) *Phrasal reconstr.*
[$_{TP}$ Jari$_1$ [$_{TP}$ T^0 [$_{vP}$ __$_1$ [$_{vP}$ v [$_{VP}$ ihaile- Merja-a]]]]]
Jari PRS.3SG (Jari.3SG) admire Merja-PAR
Agree

We will need some way to alert agreement reconstruction that T^0 requires an operation of this type. The minimalist theory available at present uses special-purpose vehicles called "probe features" for this purpose (Chomsky 2000, 2001), so let's borrow this idea and assume that there is a feature, call it [VAL] from "valuation required", which triggers the operation. We express the hypothesis in Python and run it against the test set.[5] The results are shown in Figure 11.

[5] The assumptions stated in the main text are not sufficient to specify a full Python implementation. We need a mechanism for finding grammatical heads with [VAL] from the structure, since no formal algorithm can "eyeball" them. In addition, we must specify what the local configuration between T and DP is, what will happen when Agree occurs, what happens if there are several arguments at the edge of vP, how DP adverbials are separated from arguments, and so on. Some of these issues will be examined below.

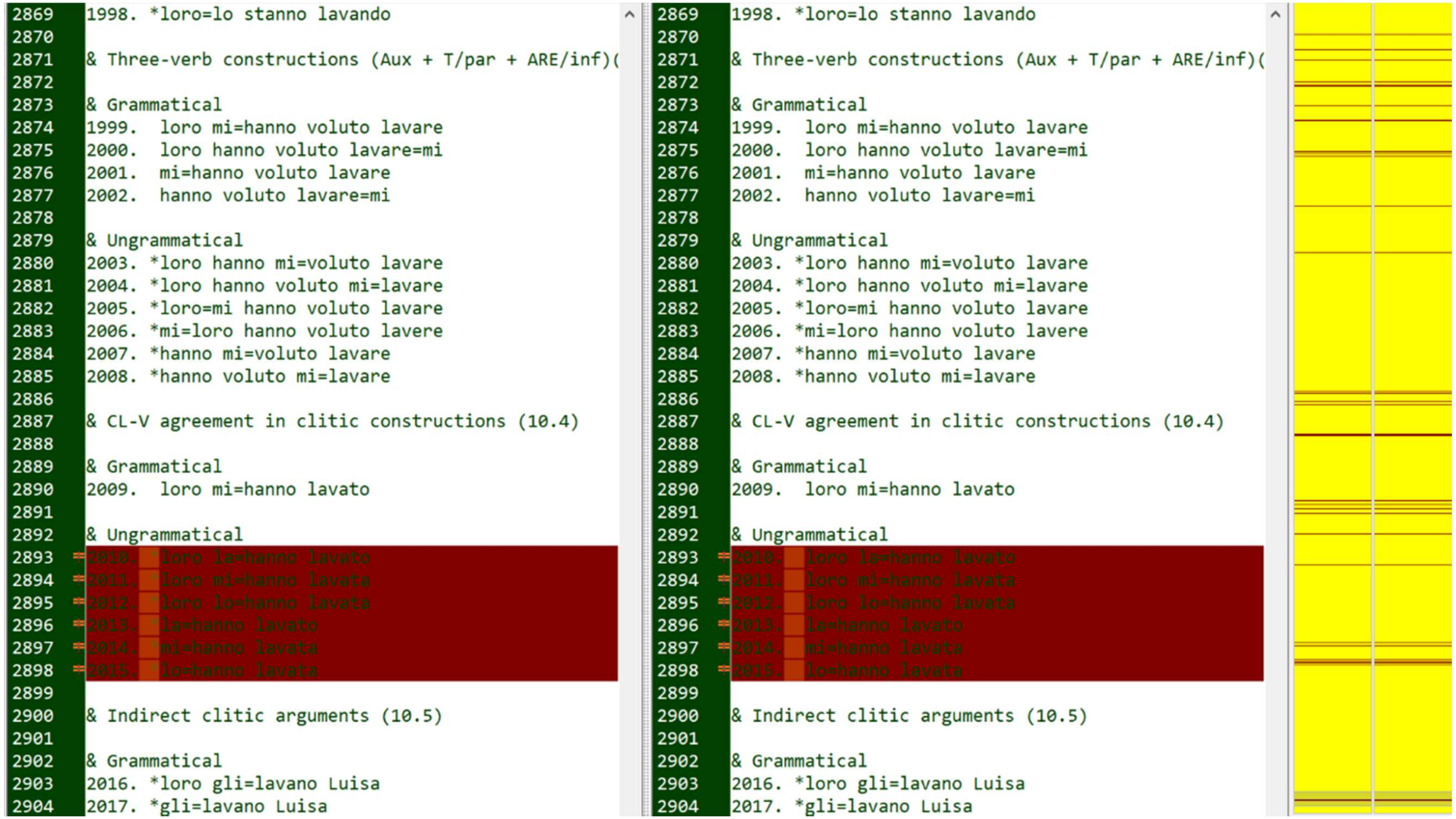

Figure 11: *Calculated results from the first trial simulation*

Several errors emerged (compare Figure 11 with Figure 2). One clear cluster is Italian clitic agreement (lines 2893–2898 in Figure 11). Properties of Italian clitic constructions indicate that there must exist an agreement dependency between a head and its specifier (i.e. clitic–participle agreement), which (3) ruled out. I return to this below; for the time being we consider Finnish. By examining the discrepancies in the Finnish datapoints we find, first, that the model is unable to handle *pro*-drop constructions. The subject pronoun can remain null in Finnish if it is not in the third person (4a–b) (Vainikka 1989, Vainikka & Levy 1999, Holmberg 2005, Holmberg & Sheehan 2010).

(4) a. *Siivo-sin* *koko* *päivän.*
clearn-PST.1SG all day
'I cleaned all day.'
b. **Siivo-si* *koko* *päivän.*
clean-PST.3SG all day
'He cleaned all day.'

First, though, we must note that (3) becomes vacuous under these circumstances. There is no overt argument against which the features of the verb could be checked. In addition, it is easy to image structural analyses under which the verb would wrongly agree with the DP adverbial 'whole day'. Let us fix this by reinterpreting the theory so that it only requires 'no mismatches', hence positive matching is not required. Sentences such as (4a) are now correctly judged grammatical. The problem, though, is that (4b) passes as well.

Here we have to consider an additional factor. The presence of agreement features seems to license the *pro*-drop phenomenon. If we licensed *pro*-drop everywhere without taking agreement into account, then all hypothetical English *pro*-drop sentences such as **admires Mary* would be wrongly calculated to be grammatical. To capture the contrast between (4a–b) while rejecting English *pro*-drop sentences across the board, it is usually assumed in the linguistic literature that the Finnish third person agreement features as well as English agreement are in some sense too 'weak' to license *pro*-drop. Perhaps we can think of strong agreement clusters as replacing overt pronoun subjects, in some sense. We could easily draw a distinction between weak and strong agreement markers in the lexicon. What complicates the issue is that sentences like (4b) *are* accepted in Finnish if they occur in a context where the missing third person subject can be linked with an acceptable antecedent (5).

(5) *Pekka* *väitti* *että* [*siivo-si* *koko* *päivän*].
Pekka claimed that clean-PST.3SG all day.
'Pekka claimed that he (= Pekka) cleaned all day.'

Agreement features license *pro*-drop in Finnish if and only if they are in the first or second person (4a–b) *or* they are in the third person and there is an antecedent (5). The pattern is summarized in Table 2.

	Person	
	1st, 2nd	3rd
Antecedent		
No	Yes	No
Yes	Yes	Yes

Table 2: *Licensing of pro-drop in Finnish as a function of agreement and antecedent*

At this point our analysis needs a mechanism for finding the required antecedents, but I will omit this issue here; see Brattico (2021), which follows the Holmberg–Sheehan hypothesis discussed further below. Test simulation shows that the proposed mechanism is still insufficient, however, because it accepts verb-initial sentences such as (6).

(6) **Siivo-si* *Pekka* *koko* *päivän.*
clean.PST.3SG Pekka.3SG all day
└─Agree─┘

T^0 agrees with the postverbal subject, but the clause is not grammatical. It cannot be just the weakness of the third person agreement that matters. The most likely reason for the ungrammaticality of (6) is that verb-initial clauses are generally ungrammatical in Finnish. In the generative theorizing this generalization is captured by stipulating that finite elements contain a topic-based "EPP-feature" that must be checked by a subject/topic phrase (Vainikka 1989, Vilkuna 1995, Holmberg & Nikanne 2002, Brattico 2019b, Huhmarniemi 2019). The sentence is grammatical if some phrase, typically but not necessarily the grammatical subject, is moved to the preverbal topic position. What we must assume, it seems, is that the "strong" first and second agreement features suffice to check this EPP condition, while third person features do not. There is a connection between agreement and the presence of a phrase at the local specifier position, and indeed now we recall that it was the absence of the same spec–head configuration that caused the original problem with the Italian clitic data. Not even this would be sufficient, however, since third person agreement features *do* suffice to the check the EPP if (and only if) there is the antecedent (see (5) and Table 2).

To me, the relativization of Finnish EPP to the presence of an antecedent has always represented something of an enigma. I will consider one possible avenue that I have used to obtain best results so far, although many problems remain. Following Holmberg (2005) and Holmberg and Sheehan (2010), we consider that the Finnish third person agreement features, in contrast to 1st and 2nd person, are not strong enough to check the D-feature of the finite element, which triggers antecedent search as a last resort, capturing the Finnish partial *pro*-drop signature. This mechanism was implemented into the algorithm reported in Brattico (2021). To connect the mechanism to the Finnish EPP and Italian clitic agreement data, we assume that agreement with the D-feature relies on the spec–head configuration, which then replaces the EPP requirement. Hence Finnish EPP = checking of a nominal D-feature, generating a definiteness effect instead of a topic effect (Brattico 2019b). Furthermore, we generalize agreement so that it checks elements both inside the sister and specifier of the triggering head. Test simulations showed that the Italian clitic data now comes out correctly. The Finnish EPP data follows if we assume that D-checking from the specifier position is mandatory: first and second person verbs can remain without overt subject, being strong enough to check the D-feature all by themselves, while third person verbs are too weak to check the D-feature, requiring an overt phrase at SpecTP or antecedent support. This agrees with the claim made by Holmberg and Nikanne (2002) that the Finnish topic EPP can be checked by any phrase that is "broadly referential", thus we interpret this claim as requiring that there is a D-feature inside the phrase occupying the specifier position. Third person constructions are grammatical if the D-feature is valued by antecedent recovery. The code corresponding to the revised final model is shown in Figure 12.

```
def Agree_1(self, head):

    goal, phi_features = self.Agree_1_from_sister(head)
    for phi in phi_features:
        self.value(head, goal, phi, 'sister')
    if not head.is_unvalued():
        return

    goal, phi_features = self.Agree_1_from_edge(head)
    for phi in phi_features:
        if find_unvalued_target(head, phi):
            self.value(head, goal, phi, 'edge')
```

Try to agree with elements inside sister constituent

If agreement leaves features unvalued, try to agree with specifier

Figure 12: *Python implementation of the final version*

Running the test corpus with this model shows that although the situation improves, the hypothesis fails to handle (7), which it judges wrongly as ungrammatical.

(7) *Merja-a ihaile-n.*
Merja-PAR admire-1SG
'I admire Merja.'

Because the thematic subject is covert, T^0 does not find anything from its sister and targets the partitive DP at the specifier position. Because both phi-features and the D-feature remains unvalued, spec–head agreement is triggered, which produces feature mismatch. The problem is that only nominative arguments are relevant for agreement in Finnish. In the first iterations of the model this was not an issue, because the model targeted the SpecvP position that contained either nothing or the reconstructed subject, hence we got the nominative/agreement correlation for free. But as soon as we include specifier agreement into the model, the problem of correlating agreement with nominative arguments resurfaced.

How to proceed from here is empirically unclear, but methodologically straightforward. We craft a representative test corpus that captures the whole Finnish agreement signature, possibly in conjunction with English agreement for comparison, and write a model that calculates it. The test corpus must contain at least finite clauses with and without agreement, with and without *pro*-drop, with and without antecedents, with and without preverbal phrases, and all these in many or perhaps in all possible word orders. More examples could be added, of course, if deemed relevant. This will ensure that, if nothing else, at least our hypothesis cannot be deemed "very far removed" from what counts as a valid hypothesis simply by subjective opinion.

5 Conclusions

One possible reason for the lack of progress in the less advanced sciences such as linguistics could be their stubborn use of an antiquated research method in which theories are justified by relying on some form of Augustine's divine illumination, a supreme cognitive capacity accessing the veridicality of an ambiguously formulated idea by intuition, common sense or thought experiment. Although this method can produce interesting data and theories, it is unable to produce agreement, creating an obstacle for progress. The notion of rigorous proof was considered a possible solution.

References

Brattico, Pauli. 2019a. *A computational implementation of a linear phase parser. Framework and technical documentation (Version 9.0)*. Pavia: IUSS University School for Advanced Studies.

Brattico, Pauli. 2019b. Subjects, topics and definiteness in Finnish. *Studia Linguistica* 73. 1–38. https://doi.org/10.1111/stul.12129

Brattico, Pauli. 2020. Finnish word order: Does comprehension matter? *Nordic Journal of Linguistics* 44. 38–70.

Brattico, Pauli. 2021. Null arguments and the inverse problem. *Glossa: A Journal of General Linguistics* 6. 1–29. https://doi.org/10.1017/S0332586520000098

Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton. https://doi.org/10.1515/9783112316009

Chomsky, Noam. 1964. *Current issues in linguistic theory*. Berlin: Mouton. https://doi.org/10.1515/9783110867565

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA.: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89–156. Cambridge, MA.: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–37. Cambridge, MA.: MIT Press.

Chomsky, Noam. 2008. On phases. In Carlos Otero, Robert Freidin & Maria-Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, 133–66. Cambridge, MA.: MIT Press. https://doi.org/10.7551/mitpress/9780262062787.001.0001

Dyson, Freeman J. 1949. The radiation theories of Tomonaga, Schwinger, and Feynman. *Physical Review* 75. 1–32. https://doi.org/10.1103/PhysRev.75.486

Holmberg, Anders. 2005. Is there a little pro? Evidence from Finnish. *Linguistic Inquiry* 36. 533–64. https://doi.org/10.1162/002438905774464322_

Holmberg, Anders & Nikanne, Urpo. 2002. Expletives, subjects and topics in Finnish. In Peter Svenonius (ed.), *Subjects, expletives, and the EPP*, 71–106. Oxford: Oxford University Press.

Holmberg, Anders & Sheehan, Michelle. 2010. Control into finite clauses in partial null-subject languages. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan (eds.), *Parametric variation: Null subjects in minimalist theory*, 125–52. Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9780511770784.004

Huhmarniemi, Saara. 2019. The movement to SpecFinP in Finnish. *Acta Linguistica Academica* 66. 85–113. https://doi.org/10.1556/2062.2019.66.1.4

Koskinen, Päivi. 1998. *Features and categories: Non-finite constructions in Finnish*. PhD dissertation, University of Toronto.

Manninen, Satu. 2003. *Small phrase layers: A study of Finnish manner adverbials*. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/la.65

Phillips, Colin. 1996. *Order and structure*. PhD dissertation, MIT.

Vainikka, Anne. 1989. *Deriving syntactic representations in Finnish*. PhD dissertation, University of Massachusetts Amherst.

Vainikka, Anne & Levy, Yonata. 1999. Empty subjects in Finnish and Hebrew. *Natural Language & Linguistic Theory* 17. 613–71. https://doi.org/10.1023/a:1006225032592

Vilkuna, Maria. 1995. Discourse configurationality in Finnish. In É. Kiss Katalin (ed.), *Discourse Configurational Languages*, 244–68. Oxford: Oxford University Press.

Pauli Brattico
Independent Researcher
pauli.j.brattico@gmail.com