Orsolya Takács: Changing trends in psycholinguistic parsing theory
Argumentum 9 (2013), 353-362
Debreceni Egyetemi Kiadó

Tanulmány

Orsolya Takács

Changing trends in psycholinguistic parsing theory*

Abstract

The present paper describes four theoretical frameworks in human syntactic parsing research that represent milestones in the evolution of comprehension studies. The aim of this concise overview is to highlight certain tendencies of change that define the future of comprehension and parsing research; and also, to report about novel directions that may take over the place of the earlier dominant models.
Keywords: sentence comprehension, parsing, heuristic strategies, Garden-Path Theory, Constraint-Satisfaction Models, shallow processing, depth of processing, Good-Enough Approach

1 Introduction

This paper aims to give an outline of the changes that took place in psycholinguistic parsing research[1] through the course of the short history of the discipline by describing two of the most influential parsing models from the past and two groups of novel theories that may represent directions to be followed in the future. The brief overview below is not by any means intended as an exhaustive account on all of the parsing models having been proposed up to the present; it is only a demonstration of the paradigm shifts that took place in the field up to this point.

2 Bever's heuristics

The study of human syntactic parsing, that is, research on how structural relations between the recognized words are computed in the mind, began just as early as that of any other subprocesses of language comprehension: in the 1950s and 60s. One of the earliest accounts on parsing hypothesized the operation of *simple heuristic strategies* behind sentence comprehension.[2]

* Work on the present paper was supported by the TÁMOP 4.2.2/B-10/1-2010-0024 project. The project is implemented through the New Hungary Development Plan, co-financed by the European Social Fund and the European Regional Development Fund.

1 It is important to emphasize that this discussion is about models of psycholinguistic parsing only; parsing used in programming languages and computational linguistics is a significantly different and more mechanical process. Still, these processes are not completely irrelevant here either, since human parsing has been looked upon as a similarly rule-governed and rigidly predetermined process for a long time. This may be one of the hindering factors in the possibly more rapid development of psycholinguistic parsing theory.

2 A heuristic strategy is to be understood as, in some sense, the opposite of an adequate syntactic algorithm: when heuris-

Orsolya Takács: Changing trends in psycholinguistic parsing theory
Argumentum 9 (2013), 353-362
Debreceni Egyetemi Kiadó

Bever (1970) suggested that the human parser operates with the help of "perceptual strategies" like the ones below:

> Strategy A: Sequence together any sequence X…Y, in which the members could be related by primary internal structural relations, 'actor, action, object…modifier'. (290)
> Strategy B: The first N…V…(N)…clause (isolated by Strategy A) is the main clause, unless the verb is marked as subordinate. (294)

In other words, speakers of English normally interpret any noun-verb-noun string as agent-action-patient, and this string of words is taken to be the main clause if the sentence is complex (unless the verb is subordinate). According to this, the sentence *The poison killed the cockroach*, for instance, is parsed with a simple application of the actor-action-object pattern: the first noun phrase *the poison* is identified as actor, the verb *kill* as action, and the second noun phrase *the cockroach* as object. Bever came up with a whole inventory of such strategies (more than a dozen), but it is only Strategy A, often dubbed as the NVN strategy, that has remained relevant and that is still referred to in recent studies. Parsing theories attributing a significant role to the thematic role structure of words and to the mental lexicon in general are particularly likely to agree with Bever on important points.

Bever's theory was, in some sense, a cognitivist one: he aimed at studying how the properties of linguistic structure in speech behaviour reflect certain general cognitive laws, and did so to offer an alternative approach to the then predominant generative theories in 'mental grammar' research (e.g. Miller's Derivational Theory of Complexity, 1962). These theories attempted to give a formal description of human linguistic competence by mapping abstract linguistic structures (i.e. Chomsky's transformational generative grammar) onto speech behaviour (cf. Chomsky's view on competence and performance; 1965). The proposition that languages are not that different from one another after all, that there may be some basic grammatical properties common in all human languages (ibid.), and that grammar may somehow be genetically encoded in the human mind (cf. universal grammar; Chomsky 1965, 1980), was a revolutionary idea, which was probably the most important trigger for a host of psycholinguistic parsing theories to appear at that time.[3]

Bever *did* use the terminology of transformational generative grammar (e.g. "underlying structure", "surface structure"; 1970: 289) for the description of phenomena observed in his "click-location" experiments performed in cooperation with Fodor and later with Garrett (Fodor & Bever 1965, Garett et al. 1966).[4] However, in opposition with Miller, his point was that this abstract grammar may not be mapped directly onto speech production, and it is "perceptual strategies" (i.e. heuristics) that help us identify the deep structure of sentences rather than the direct (and perhaps conscious) application of intricate syntactic algorithms. He even noted that these heuristic strategies only capture generalizations that are not necessarily true in each case, but they hold for most of the cases. "Semantic strategies" were also mentioned in

---

tics are applied, the interpretation of the sentence develops "on the way", time constraints and fast processing are more emphatic than infallible precision, and exceptions to which the "rule" does not apply are likely to occur.

[3] However, linguists' efforts to create a syntactic theory applicable to any construction in any human language have not been successful up to the present, nor the attempts to propose a theory of 'mental grammar' that verifiably describes the way the human brain in fact uses syntax in production and comprehension processes. Furthermore, there is growing evidence that only a restricted application of linguistic grammars is possible in comprehension theory, and lexical-semantic (as well as some other) factors have an important role in how "mental grammar" works (cf. e.g. MacDonald et al. 1994).

[4] The basic layout of these experiments was the following: subjects listened to sentences, during the course of which they always heard a non-speech interruption at some point. Based on their indications about these "clicks", the authors concluded that subjects always reported them to have occurred closer to deep structure boundaries than they actually and objectively did. The idea of *heuristic strategies* was derived from this curious relation between surface and deep structure boundaries in perception.

Orsolya Takács: Changing trends in psycholinguistic parsing theory
Argumentum 9 (2013), 353-362
Debreceni Egyetemi Kiadó

his paper, referring to certain kinds of semantic constraints (cf. plausibility) and frequency information, both of which have been studied extensively since (cf. constraint-satisfaction models of parsing).

Bever's conception was rejected back then, because it was considered to be too simple for the handling of a number of complicated English structures, and its explanatory force was restricted to only a limited number of sentences. Even though those structures may be the ones most often used in spontaneous everyday communication, still this account was not a complete theory of sentence comprehension. Nevertheless, as empirical research in psycholinguistics is helped more and more by the advances of technology (e.g. rapidly developing neurological imaging techniques, eye-tracking applications, reading and reaction time studies), some very recent, intriguing findings seem to prove a number of other theories wrong, and the possible viability of heuristic strategies in real-life speech production and comprehension is often reconsidered.

3 The garden-path theory

Not long after Bever, Frazier proposed another, much more influential model of parsing, or rather the model gradually emerged from a series of papers (1978, 1979, 1987). This account is most often referred to as the *garden-path theory*, and it is still the most widely accepted model of syntactic parsing since. Sentences like (1) or (2) probably sound familiar to anyone who has ever read or learnt about psycholinguistics.

(1) *The horse raced past the barn fell.*
(2) *The girl knew the answer was correct.*

These are commonly termed as "garden-path" sentences, named so because temporarily ambiguous sentences like these contain syntactic structures that "lead the comprehender up the garden path" causing them to initially misanalyse the sentence, and then forcing them to reanalyze. Reduced relatives (also referred to in the literature as relative clause attachment ambiguities or main clause/relative clause ambiguities) and object/object complement ambiguities similar to (1) and (2), respectively, have been much studied constructions in parsing research since Frazier started work on them. In the case of (1), comprehenders tend to choose the main clause interpretation (Figure 1 a) of the verb *raced* first, and it is only when they encounter the verb *fell* that they notice the error and reanalyze the sentence with the reduced relative interpretation (Figure 1 b).

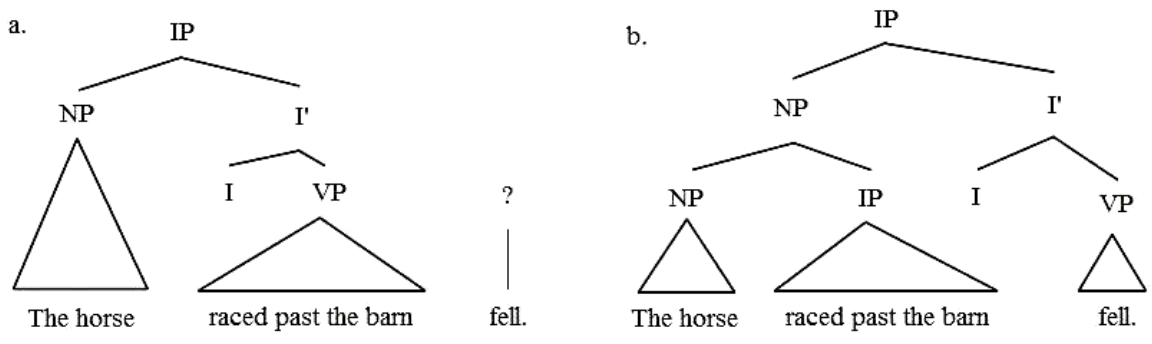

Figure 1: The phrase structure of the classical reduced relative ambiguity

Orsolya Takács: Changing trends in psycholinguistic parsing theory
Argumentum 9 (2013), 353-362
Debreceni Egyetemi Kiadó

Similarly, the parsing of (2) is difficult, because the reader first takes the noun phrase *the answer* to be the object of the verb phrase headed by *knew* (Figure 2 a), and realizes only later that it is the subject of a new complement clause. The analysis in which *the answer* is the object of *knew* becomes impossible only when the word *was* is encountered (Figure 2 b).

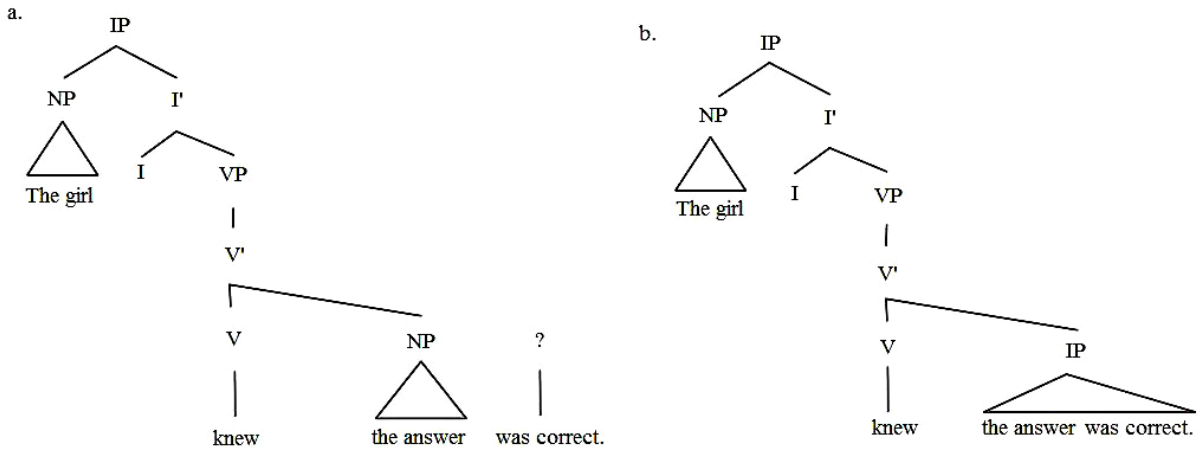

Figure 2: The phrase structure of an object versus sentential complement ambiguity

As an explanation for phenomena like these, Frazier suggested that the human parser is strictly modular, and initially, it constructs a single phrase structure for the heard or read sentence based solely on the grammatical categories of the words in the lexical string. Neither semantic nor lexical knowledge is available in this early stage of parsing, and these types of information are only used during reanalysis, if the parser is "garden-pathed". When the chain of grammatical categories is compatible with more than one analysis, two principles guide the parser in the resolution of the syntactic ambiguity: (a) Minimal Attachment and (b) Late Closure.

The *Minimal Attachment* principle says that the parser attaches incoming material to the phrase structure tree using the fewest possible number of new constituents, preventing the postulation of unnecessary nodes. This causes the misanalysis of both (1) and (2), since the garden-path interpretation requires one fewer node than the correct analysis, only this simpler analysis is blocked before the sentence ends.

The other principle, *Late Closure*, establishes that new material is attached to the node currently being processed rather than to a former constituent or to a new one. A good example for the operation of this principle is the sentence *When Mary was knitting the socks fell to the floor*. Here the noun phrase *the socks* is initially taken to be the direct object of the verb phrase *was knitting* instead of being correctly interpreted as the subject of a new clause.

In other words, the human syntactic processor always chooses the structurally simpler analysis, with no unnecessary nodes postulated, and by skipping back and forth between nodes only if it is absolutely necessary. This preference for simplicity is often termed as "preference for low attachment" or "preference for right-branching constructions" as well. These expressions refer to the spatial position of the new constituent in the phrase structure tree.

The earliest version of the garden-path theory was already contained in Fodor and Frazier's "Sausage Machine" model (Frazier 1978). In this, the authors proposed that the syntac-

Orsolya Takács: Changing trends in psycholinguistic parsing theory
Argumentum 9 (2013), 353-362
Debreceni Egyetemi Kiadó

tic representation of a sentence is formed in two steps: first, lexical or phrasal nodes are assigned to six-word substrings of the incoming material,[5] then the parser combines these phrases into a complete sentence structure by adding higher nodes to link them together. They hypothesized two substructures for the parser to carry out these steps, which were named the "Preliminary Phrase Packager" (or the "Sausage Machine") and the "Sentence Structure Supervisor". The PPP was considered to be more or less insensitive to well-formedness rules and to be "short-sighted" (i.e. seeing only about six words at a time). The SSS, on the other hand, was thought to survey the whole of the sentence and to keep track of long-distance dependencies as the structure is computed. The Minimal Attachment principle, proposed here for the first time, was the general and only principle guiding the overall structuring of sentences. The other principle, Late Closure was added to the theory only later (see Frazier 1979, 1987). An important and interesting detail about the "Sausage Machine" is that the authors admittedly based their account on a critique of the ATN model (Woods 1970) and the heuristic strategies hypothesized by Bever (1970) and Kimball (1973). Of the three, they could agree with Kimball's theory the most, and some component parts of his seven principles were adopted by them in a modified format: Minimal Attachment is an adaptation of Right Association[6] and Late Closure is a modification of Closure[7].

The final version of the garden-path theory, in so far as the theory can be considered as finished, was developed by Frazier alone in her subsequent papers (1979, 1987). The most important alterations compared to the "Sausage Machine" are the omission of the hypothesized PPP and SSS modules, the addition of the Late Closure principle, and a specification of the types of information used during initial parsing processes and through the course of reanalysis. She also discusses the so-called "filler-gap" dependencies (or long-distance dependencies; e.g. *What did John eat?*) and proposes a "recent filler strategy" for the handling of these. This strategy, however, is less known, accepted and applied than the other two principles.

4 Constraint-satisfaction models

While former parsing models were only tested experimentally and by means of grammaticality judgement tests, Frazier's model has been comprehensively tested with technologically more complex and developed methods like eye tracking, ERP studies[8] and reading time experiments. The new technology available helped to make some well-established observations on how syntactic ambiguities are resolved. The results of these studies have been inconclusive, but part of the findings led some researchers to think that the mental lexicon and the syntactic parser were not completely separate modules in the mind.

A 1994 comprehensive article of MacDonald and colleagues titled *The lexical nature of syntactic ambiguity resolution* is one of the most influential and still often-cited papers suggesting that semantic and contextual effects have a much more significant role in initial parsing processes already, and that phrase structure rules and proper syntactic algorithms are

---

5 This number has to do with the often cited fact that the capacity of the short-term (or working) memory is as much as 7±2 digits.

6 "Terminal symbols optimally associate to the lowest nonterminal node." (Kimball 1973: 24)

7 "A phrase is closed as soon as possible, i.e., unless the next node parsed is an immediate constituent of that phrase." (ibid., p. 36)

8 Event related potentials. The measurement of brain responses to specific linguistic tasks with electroencephalography (EEG).

Orsolya Takács: Changing trends in psycholinguistic parsing theory
Argumentum 9 (2013), 353-362
Debreceni Egyetemi Kiadó

not directly applicable in real-life human sentence processing; most of the syntactic information is more likely to be stored in the mental lexicon instead.

The goup of parsing accounts that were created in the spirit of this lexicalist "ideology" is that of the non-modular (or interactive) models, the majority of which belong to the category of the so-called *constraint-based (or evidential) model*. Their most important claim is that all the possible information sources (lexical, contextual and syntactic) are immediately available for the processor in the initial parsing stage already, and parsing is not serial or modular as the garden-path theory proposed. Multiple analyses are activated (not only a single parse) and all the possible interpretations of a sentence compete for selection. The evaluation of the various analyzes is continuous, and the choice between them is influenced by how much support one or the other analysis gets from the available knowledge sources: lexical constraints (like semantic plausibility, verb subcategorization preferences etc.) narrow down the number of possible interpretations. When one analysis receives more support than the others, processing is unhindered. Difficulty occurs only when two analyses get equal support.

Constraint-based models appeared in the early 1990s, and the most important claim against them was that they were too vague, underspecified and unfalsifiable. The identification of all possibly existing constraint and the way of combining and weighting them seemed (and still seems) to be a task impossible to solve through computational linguistic modelling, and thus, they are untestable. The best-known of these models that has already been computationally implemented is the *Competition-Integration Model* by Spivey-Knowlton, Tanenhaus and McRae (Spivey-Knowlton 1994, Spivey & Tanenhaus 1998, McRae et al. 1998).

5 Shallow processing accounts

In the past three or four decades, the garden-path theory has been very dominant in psycholinguistic parsing research as a theoretical basis for empirical experiments, and for a long time, only constraint-based models could be considered as essentially different, but viable alternatives to it. From the beginning of the 21st century, however, a new theoretical direction has begun to emerge from certain findings about garden-path sentences and other types of "difficult-to-process" structures. This trend is rather novel as yet, without finished and sufficiently elaborated theories. Nevertheless, there is an emerging new group of sentence processing accounts that may be called the *shallow processing theories*.

In their 2002 article, Sanford and Sturt pointed out that the representations and interpretations that are built during sentence comprehension, are often "underspecified" in one respect or another. This means that a basic assumption of every current theory of sentence processing, namely that the syntactic representations created by comprehenders are always complete and correct, has been questioned. According to the authors, such fully specified representations are often not desirable anyway for the tasks comprehenders need to perform in real-life, online communication. The sentences below are examples from Sanford and Sturt, provided as evidence for various cases of "shallow" processing in language comprehension.

(3) *A: 'can we kindly hook up…uh…engine E2 to the boxcar at Elmira'*
B: 'okay'
A: 'and send it to Corning as soon as possible please'
B: 'okay'

(4) *How many animals of each sort did Moses put on the Ark?*
(5) *No head injury is too trivial to be ignored.*

The first item, (3), is a dialogue about train scheduling, coming from the TRAINS corpus of the University of Rochester. It illustrates that the reference of pronouns is often vague in everyday communication (the exact referent of *it* being unspecified here), but this does not seem to influence our general understanding of the dialogue. The second example is often referred to as the "Moses illusion", and it is relevant because of people's inability to detect a semantic anomaly when the referents of two expressions are very similar, and the focus of the discourse direct their attention to somewhere else in the sentence (i.e. people do not usually register that Noah built the Ark and not Moses, and they answer 'two' to the question). Finally, example (5) is a demonstration of how inferences from general knowledge can predetermine the interpretation of a sentence at an early stage of processing. It is almost impossible to recover from the misinterpretation. We take the sentence to mean 'Whatever minor a head injury seems to be, it still should be treated', while the meaning is just the opposite: 'No matter how trivial a head injury appears to be, it still should be ignored.'

Interestingly, the concept of *shallow processing* originally comes from computational linguistics (there, it is a simple identification of non-overlapping groups of words, as opposed to a full-fledged and detailed syntactic parse), but it may be much more relevant to human parsing processes than it is currently assumed. Nevertheless, the raison d'etre of the concept is basically the same in both cases: an economical exploitation of the resources at hand.

Although the evidence the authors listed in this article is largely of semantic, contextual and pragmatic nature, the issue they raised seemed very much relevant and intriguing from the point of syntactic comprehension research as well. Another article came out in the same year from Ferreira, Bailey and Ferrano (Ferreira et al. 2002) with a similar hypothesis, but written in a more parsing-centered and more psycholinguistically oriented manner. The fundamental idea was, again, that incorrect and incomplete semantic and syntactic representations are not rare in online human sentence comprehension. However, the authors in this case also provided an outline and description of the traditional parsing models (the garden-path theory and constraint-satisfaction models) and compared the claims of those with some of their own empirical findings. Part of the research Ferreira and colleagues refer to during the elaboration of their theory overlap with examples from the above mentioned Sanford–Sturt article, but Ferreira's own experiments and research material also contribute to the justification of the hypotheses and questions raised by the two authors earlier.[9]

6 Tendencies of change

Taking a look at all the models described above together, some characteristic tendencies of change in parsing theory are observable.

The changing status of the grammatical and the semantic components in human sentence comprehension is one of the key issues when trying to grasp the nature of changes across the above described models. A notion central to this topic is *modularity*. Traditionally, linguistic processing was thought to consist of distinct modules that involve different types of information and processes. According to Fodor's concept of modularity (1983), each module responds only to the type of information it is responsible for, and there is no communication

[9] For details about Ferreira's theory, see e.g. Ferreira et al. 2002, Ferreira & Patson 2007.

between the modules. The most important aspect of this conception from the perspective of parsing theory is that the lexicon (handling word-level processing) and the syntactic parser (computing grammatical representations) do not interact in any way – lexical and syntactic processing happen independently from each other. Two-stage parsing theories, most importantly Frazier's garden-path theory (1979), applied this notion of modularity. From the 1980s on, however, most of the theories hypothesized the immediate availability of all types of information in the initial stages of parsing already, with no separate modules differentiated (cf. constraint-satisfaction models, shallow processing accounts).

Furthermore, there seems to be another important factor that triggered tendencies of change through the course of parsing research, and it is the question of how detailed syntactic analysis is. From this respect, the garden-path theory is in opposition with all the other models described in this article, since Bever's theory, constraint-satisfaction models and shallow processing accounts all suggested that the human parser does not always compute a detailed syntactic representation for sentences, and the nature of the comprehension process is often semantic rather than syntactic. Although, Frazier's model is still the most influential parsing model, recent experiments referred to by MacDonald (1994), Sanford & Sturt (2002) and Ferreira et al. (2002, 2007) make us reconsider the assumption that the syntactic structure we compute when processing a sentence is necessarily complete and correct in each of the cases.

A further important aspect of human communication justifying such considerations is the role of *memory and time constraints in real-life comprehension*. For a long time, these constraints have not really been considered as important in the development of processing theories due to the hypothesized priority of grammar in comprehension and production, and because of a strict theoretical separation of competence and performance (cf. Chomsky 1965). However, the studies that the proponents of constraint-satisfaction (see e.g. (Spivey-Knowlton 1994, Spivey & Tanenhaus 1998, McRae et al. 1998) and shallow processing accounts cite (see e.g. Sanford & Sturt 2002, Ferreira & Patson 2007), have proven that "the finite set of rules" that may be termed as a set of syntactic algorithms cannot always be applied when we use language in real-life contexts, due to lack of time or memory capacity. As a consequence, the former group of theories rejects grammatical algorithms wholly and completely; while shallow processing theories propose a framework in which language understanding is fundamentally not a syntax-driven process, but during this process, syntactic algorithms do operate in certain conditions.

7 Conclusion

The past seventy years of psycholinguistic research saw a number of different theories of parsing. This article showed just a sample of those, but it is nevertheless obvious from this limited overview already that there have been considerable changes in how researchers think about the human sentence comprehension process. The emphasis seems to be switching from syntax to semantics and contextual factors, and the strict modularity of the human parser is being rejected with silent agreement due to recent findings. It remains to be seen whether the new theories described here will overturn the garden-path theory, or some other models will emerge instead, but the diversity and novelty of directions taken seems to signal another paradigm shift in parsing studies.

References

Bever, T. G. (1970): The cognitive basis for linguistic structures. In: Hayes, R. (ed.): *Cognition and Language Development.* New York, Wiley & Sons, Inc. 279–362.

Chomsky, N. (1965): *Aspects of the Theory of Syntax.* Cambridge, MA: MIT Press.

Chomsky, N. (1980): *Rules and Representations.* New York: Columbia University Press.

Ferreira, F. & Patson, N. D. (2007): The 'Good Enough' Approach to Language Comprehension. *Language and Linguistics Compass* 1/1–2, 71–83.

Ferreira, F., Bailey, K. G. D. & Ferrano, V. (2002): Good-enough representations in language comprehension. *Current Directions in Psychological Science* 11(1), 11–15.

Fodor, J. A. (1983): *Modularity of mind.* Cambridge, MA: MIT Press.

Fodor, J. A. & Bever, T. G. (1965): The psychological reality of linguistic segments. *Journal of Verbal Learning & Verbal Behavior* 4, 414–420.

Frazier, L. & Fodor J. D. (1978): The sausage machine: a new two-stage parsing model. *Cognition* 6, 291–325.

Frazier, L. (1979): *On comprehending sentences: Syntactic parsing strategies.* PhD Dissertation, Indiana University Linguistics Club, University of Connecticut.

Frazier, L. (1987): Sentence processing: A tutorial review. In: Coltheart, M. (ed.): *Attention and performance XII: The Psychology of Reading.* Lawrence Erlbaum Associates.

Garett, M., Bever T. G. & Fodor J. D. (1966): The active use of grammar in speech perception. *Perception and Psychophysics* 1, 30–32.

Kaplan, R. & Bresnan, J. (1982): Lexical-Functional Grammar: A Formal System for Grammatical Representation. In: Bresnan, J. (ed.): *The Mental Representation of Grammatical Relations, Chapter 4.* Cambridge, MA: MIT Press, 173–281.

Kimball, J. (1973): Seven principles of surface structure parsing in natural language. *Cognition* 2:1, 15–47.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994): The lexical nature of syntactic ambiguity resolution. *Psychological Review* 101(4), 676–703.

McRae, K., Spivey-Knowlton, M. J. & Tanenhaus, M. K. (1998): Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language* 38, 283–312.

Miller, G. A. (1962): Some psychological studies of grammar. *American Psychologist* 17, 748–762.

Sanford, A. J. & Sturt, P. (2002): Depth of processing in language comprehension: not noticing the evidence. *Trends in Cognitive Sciences* 6, 382–386.

Spivey, M. J. & Tanenhaus, M. K. (1998): Syntactic ambiguity resolution in discourse: Modeling the effects of referential context and lexical frequency. *Journal of Experimental Psychology: Learning, Memory and Cognition* 24, 1521–1543.

Spivey-Knowlton, M. (1994): Quantitative predictions from a constraint-based theory of syntactic ambiguity resolution. In: Mozer, M. C., Touretzky, D. S. & Smolensky, P. (eds.): *Proceedings of the 1993 connectionist models summer school*, Hillsdale, Lawrence Erlbaum Associates, 130–137.

Woods, W. (1970): Transition Network Grammar for natural language analysis. *Communications of the ACM* 13, 591–606.

Orsolya Takács
University of Debrecen
Institute of English and American Studies
H-4010 Debrecen
Pf. 73
lola.red@gmail.com