DRAFT, Sept. 6, 2016. To appear in: de Almeida, R. G. & Gleitman, L. (Eds.) (forthcoming). *Minds on Language and Thought.* Oxford University Press.

***The unity of consciousness and the consciousness of unity*[1]**

T.G.Bever

University of Arizona.

Truth is stranger than fiction….because fiction is obliged to stick to the possibilities. Truth isn't

Mark Twain, "Following the Equator"

ABSTRACT

Every language-learning child eventually automatically segments the organization of word sequences into natural units of composition – a phenomenon reflected in the classic studies of where adults form natural superlexical units as revealed by where they perceive "clicks" during sentence comprehension. Recent research shows that within the natural units, there is a temporal disconnect between listener's awareness of the sound and representation of other levels such as the lexical sequence and the meaning of utterances: in critical cases a compressed or absent word at a point early in a sequence is unintelligible until *later* acoustic information; yet listeners think they perceived the earlier sounds and their interpretation *as they were heard*. There are examples of conversational speech directed to both adults and children that have this property (at least for adult listeners). This discovery has several implications: (a) the "poverty of the stimulus learning problem" emphasized by Fodor for cognitive science in general, is far graver for language than usually supposed. (b) Our conscious unified experience of language as we hear it and simultaneously interpret it is in part reconstructed in time-suspended "psychological moments"; (c) The serial domain within which such integration occurs may be the structurally defined "phase" from current minimalist theories, which unifies the serial percept with structural assignment and meanings; (d) levels of language processing overlap in a "computational fractal": each level involves the same kind of analysis-by-synthesis interaction of associative-serial and structure dependent processes.

KEY WORDS (no particular order)

poverty of the stimulus, perceptual units, conversational speech, phase, consciousness, click location, structure dependence, intonation and phrasing, deep structure, analysis-by-synthesis, the LAST comprehension model, "good enough" processing, computational fractal, opera.

1. A sentence is like a (miniature) opera.

Music is often analyzed in relation to language to give perspective on the structural and formal aspects of language. But even the simplest sentence surpasses what music can tell us about it. A sentence in everyday use combines a stream of sound, with rhythm and pitch variations, with memorized units of meaning, an organizing structure that recombines those meaning units into a transcendental unified meaning that includes informational representations, general connotations, and specific pragmatic implications unique to the conversational context.

In other words, each sentence is a miniature opera of nature.

Children grow up surrounded by one opera after another, and miraculously learn to create their own. This is achieved in the context of experiencing only a small number of fully grammatical sentences, many ungrammatical ones, and very little specific feedback on their mistakes. This situation is generally referred to as "the poverty of the stimulus", which is the basis for the argument that much of linguistic structure must be innately prefigured (Chomsky, 1959,1965,1975,1980). Fodor (1981) broadened the implications of this argument beyond language into cognition in general: "The [argument from the poverty of the stimulus] is *the* existence proof for the possibility of a cognitive science" (p. 258). Nonetheless, the flagship case of the argument continues to be the speed with which children learn language in erratic environments with variable feedback.

In this paper, I focus on one of the major components of what the child has to discover in learning to understand and then produce language – the perception and comprehension of natural units of composition in the serial string. Interestingly, this problem exists on virtually any grammatical theory, from taxonomic phrase structure all the way up to today's Minimalism. Every view of what language is, going back centuries has some notion of serial hierarchical phrasing as a fundamental component. In phrase structure grammars, describing the phrase is the direct goal and what the child must discover; in generative theories that utilize "structure dependence" the child must discover the phrase in order to have access to the structure. In the next sections, I trace some research on how major language units are perceived, over the past decades, and then turn to the implications of recent studies of the acoustics of normal conversation, which show how deep and puzzling the problem of the poverty of the stimulus really is. The processing of normal conversation reveals a disconnect between the listener's representation of the sound and meaning of utterances. In critical cases it is possible to show that compressed or absent words are unintelligible until the listener hears *later* acoustic information Yet the listener perceives the acoustic presentation of the words as simultaneous with the comprehension of it. This is an instance of *creating* a conscious representation retrospectively.

I draw a number of morals from such facts in language processing, notably: the "poverty of the stimulus problem" is far graver than usually supposed – although the words in some child-directed speech are carefully pronounced, many are not. And children are also surrounded by the same kind of garbled and cue-poor instances from adult speech; this means that structure dependence must guide ongoing comprehension processes of externalized serial input, not used only to decide about the abstract structure of one's language during learning; every level of language experience involves some encoding: this supports the notion that ongoing processing occurs in a set of simultaneous parallel processes in a “computational fractal”, that is, each level involves the interaction of associative-serial and structure dependent processes; thus, our conscious experience of language is in part reconstructive in temporarily time-free “psychological moments” – so language comprehension processes move forwards *and backwards*, even though the phenomenal experience is that it moves only forward.; this reconstructive analysis of our conscious experience of language may be typical of other modalities of our experience.

This leads us to distinguish the computational problem of language acquisition from the acoustic input problem. The computational problem concerns how children generalize in the right way from scant examples of complete well formed sentences with clearly presented words, how they alight on the right kind of structure dependent hypotheses. The acoustic input problem is that children (and adults) are often not presented with clear word-by-word inputs to learn and understand from. Rather children must have already solved a large part of the computational problem in order to resolve the acoustic input problem. This magnifies what we must assume is available to the child at a very young age, and geometrically complicates any attempts to model acquisition with statistical models unadorned by massive and keen prior structures and expectations.

2. Where is the unit of language processing?

Psychology as a field often depends on resurgent methodology and continually mysterious phenomena: One of the most enduring methods and mysteries is the systematic mislocation of “clicks” presented during auditory presentation of sentences towards “phrasal” boundaries of some kind. The use of click-mislocation was pioneered by Ladefoged and Broadbent (1960), as a way of showing on-line segmentation of syllables. Its utility for exploring on-line complexity and the effect of “phrase” boundaries was initially explored by Garrett (1964. Fodor and Bever (1965) demonstrated the general role of relative depth of surface “phrase” breaks in determining the likelihood of click mislocation to or towards them; Garrett, Bever and Fodor (1966) showed that the mislocation was not due to local intonational cues, but to the “phrasal” structure that listeners impose on what they are hearing. (For a contemporary demonstration of brain spectral activity corresponding to phrase construction without benefit of intonational or statistical cues, see Ding, et al (2016) and discussion below).

A revealing aspect relevant for today’s discussions is the fact that the citation of the two original click location papers has experienced a “U shaped function” with almost as many citations in the last five years as in the first five, and less than a third of that rate in the intervening years . This reflects the rediscovery of questions about what the “real” unit of ongoing language processing is.

Later studies attempted further to define what perceptual and comprehension units are revealed by click mislocations, “deep” structure units (Bever, Lackner, and Kirk (1969)) or ”major” surface phrases (Chapin. Smith, and Abrahamson (1972). Many click location studies required subjects to write out the sentence and indicate the click location – this invited the interpretation that the click mislocation effect was not perceptual at all, but some form of response strategy related to recapitulating the sentence. Bever et al (1973) explored this by having listeners mark the click location within a window in the text written out and presented right after hearing the stimulus. In critical cases there was no auditory click at all: to make it plausible that there was a click, the loudness of the actual clicks was varied. When a click was present, the usual effect of a major phrase boundary occurred: when there was no click, subjects' guesses were *not* systematically placed in the major phrase boundary. Using a different method, Dalyrymple-Alford (1976) confirmed that click mislocation is not due to a response bias.

Two approaches to the question of the processing unit have continually surfaced and resurfaced over many years: each rests on one of the two ideas dominant in centuries of cognitive science: (a) the currency of mental life is statistically determined associations; (b) mental life is organized into categorical representations. The argument started with a closer examination of the 'psychological reality of linguistic segments namely the "phrase". During the 1960s much attention was being given to the idea that "phrases" could be defined in terms of serial predictability (Johnson,1970; Osgood,1968). On this view, "phrases" are behaviorally bounded by relatively low points of serial predictability: indeed it is generally the case that phrase-final (content) words are more predictable locally than phrase initial words. So behaviors that seem to reflect categorical phrasing might actually be reflecting variation in serial predictability. However, when syntactic structures are held constant while local predictability is varied, the high serial predictable points actually attract clicks perceptually (Bever et al, 1969). So, probability governed segmentation does not account for the online perceptual formation of phrases.

Yet, the conflict between some version of association and categorical structural assignment always finds new life. The connectionist postulation of hidden units, back propagation and other statistical devices, along with the rehabilitation of Bayesian statistics, resuscitated notions of mediated associations with complex descriptive power, enabling simulation of categorical structures as their limit (e.g., Rumelhart and McClelland, 1986). In this vein, great attention is given to "feed forward" models of perception in general and sentence processing in particular: the perceptual system is constantly making predictions of what is about to occur, so that much of language interpretation is actually rolling confirmation of specific kinds of rolling perceptual expectations. In the case of language, this can occur simultaneously at various levels from the acoustic to the semantic. The expectations are arguably a blend of probabilistic and categorical features in many domains; phonological, semantic, and syntactic. Canonical demonstrations of this are effects of left->right constraints during processing: something that occurs at point a affects the perception of something later at point b.

What I will explore in the next few pages is more recent evidence that parsing is not only "forward" it is also "downward", the construction of meaning units within short epochs. The crucial demonstration of this is evidence for backward constraints: something at point *b* in a sentence determines the percept at an earlier point *a*. Most critical to this argument is that the conscious awareness is of a constant forward moving perception, not a period of blank content suddenly filled in by something that comes later. That is, we perceive sentences in "psychological moments" in which the underlying computational processing can move back and forth, or more to the point, forth and back, before 'reporting out' to conscious awareness.

3. The unity of processing units and the conscious experience of language.

Linguistic and psycholinguistic research on sentence structure and processing has implicitly assumed that the constituent words are given: that is, the syntactician's (and child's) problem is to determine the regularities that govern how the words and other syntactic units are arranged (and inflected when relevant); the psycholinguist's problem is to determine the processes that underlie how the words and units are composed together in production of sentences, mapped onto representations in comprehension of sentences, and learned in relation to their role in possible syntactic constructions. But outside of syntax classes and psycholinguistic experiments, the words in natural language are rarely clearly or fully presented – the acoustics of one word blends into another, and in many cases, large portions of a word or word sequence are actually not present at all: to borrow a term from phonology, the words are encoded together.

Some well known facts about serial encoding at the phonological level may help us understand the situation at the syntactic level. First, it is well documented that unvoiced stop consonants in English, may actually not be given any acoustic power of their own. Thus, the final consonant in the words /top/, /tot/, /toc/ may be silent or all converge on glottal stop – yet we hear them quite clearly as distinct – it is the way that the preceding vowel changes as it quickly approaches the articulated position of the consonant. If we could hear the preceding vowels drawn out in time, they would be more like /TOuP/, /TOiT/, /TOaC/: the last bit of the vowel gives the clue as to where the tongue is heading before the vowel goes silent. Yet, our conscious percept is that the consonant was actually uttered. This is an example of a "feed forward" activity, in which the material preceding the final silence or glottal stop makes a strong enough prediction about what will be "heard" so that it is actually perceived even when not in the signal itself.

But the influence of one part of a phonological sequence on another is not always "forward", it can be "backward" as well. It is well known that it is the timing of the onset of a post-consonantal vowel that communicates whether the preceding consonant is to be heard as voiced or unvoiced. Even more striking is that in some variants, the initial voiced consonant can also not be explicitly uttered: the difference between /bill,dill,gill/ can be only in the vowel transition following the initial occlusion of the vocal tract, just long enough to indicate voicing – it is the vowel transition away from the silent initial consonant (except for the voicing itself) that indicates what the preceding consonant *was*.

The moral is that at the phonological level, even when a word is uttered in isolated "citation" form, we automatically use early phonetic information to guide the conscious representation of what follows, *and conversely*.

It can be argued that at the level of individual words, this only shows that the unit of word recognition is larger than individual phonemes, for example, that listeners have prepackaged representations of entire syllables, or that different kinds of acoustic features work together in a "cohort" (see. E.g., Marlsen-Wilson and Zwitserlood, 1989). This kind of argument may be possible in principle for words and phonology, since there is a finite number of syllables used in any particular language. But as is classically argued, such proposals of memorized units become much harder to rely on at phrasal and sentential levels, since the number of different phrases and sentences is enormous, arguably infinite in the latter case. So, we might not expect both forward and backward processing interactions at these higher levels of language. But in fact, recent evidence suggests that this is the case in normal uses of language outside of the syntax classroom and laboratory.

4. Unconscious comprehension processes with backwards inferences.

The rapid and unconscious resolution of local ambiguity suggests that corresponding prospective and retrospective processes occur at the syntactic level. For this discussion , the most significant effect is the immediate role of retrospective processing that we are unaware of. If you hear a sentence like the following, in (1a,b), there can be evidence that the ambiguity of the lexically ambiguous phonetic sequence “pair/pear” creates momentary computational complexity reflected for example in decreased accuracy of a click immediately after the word (Garrett, 1964). But you are not aware of it, and have the strong impression that you assigned it the correct interpretation as you heard it. Swinney (1979) showed that both meanings of an ambiguous word facilitate an immediately following lexical decision task, even when there is a preceding disambiguating context, e.g., as in (1c,d); but a few words later, only the contextually supported meaning facilitates the task.

(1) a. The pair of doves landed on our porch.
b. The pear and apple landed on our porch.
c. The doves in a pair landed on our porch.
d. The apple and pear landed on our porch.

A series of investigations by Fernanda Ferreira and colleagues (e.g., Christianson, Hollingworth, Halliwell, & Ferreira, 2001) complements Garrett’s (1964) finding at the phrasal level. Even after a garden path in segmentation of a written sentence is corrected by later material in the sentence, listeners retain a semantic representation of the initial (incorrect) segmentation. So for example, in critical trials, they follow the sentence (2a) below with a question, to which the subjects have to report “yes” or “no” to the question in (2b)

(2) a. While Bill hunted the deer ran into the woods.

b. Did Bill hunt the deer?

c. Did the deer run into the woods?

Surprisingly, Christianson et al (2001) found that about a quarter of the responses were “yes” to (2b) following (2a). At the same time, they found that the subjects almost always answered the question in (2c) correctly: so they argued that “the reanalysis processes got as far as identifying a subject for the main clause verb, but didn’t finish up by revising the interpretation on which that same NP was once the object of the verb in the subordinate clause." What is important for my current focus is that when subjects answered (2b) correctly or not, they were quite confident in their answers. "subjects were quite poor at arriving at an interpretation licensed by the input string, yet surprisingly confident that they had correctly understood the sentences." (p. 380). Christianson et al. take this to be evidence that comprehenders construct representations that are "good enough" to contribute to ongoing comprehension, especially in normal discourse contexts (Ferreira and Henderson, 1991). Since most sentences do not have strong garden paths (especially in auditory mode), "good enough" representations are usually good enough. That is, people arrive at conceptually appropriate interpretations based on incomplete or incorrect analyses of which they are totally unaware. More recent studies support the view that subjects do in fact analyze the correct segmentation in the garden path structures on-line, even though their answers to probe questions indicate that they consciously retain the influence of the incorrect parse (Ferreira et al, 2002; Ferreira and Patson, 2007; Ferreira et al 2009; Slattery et al, 2013).

A classic line of research on backward influences on processing started with the studies by Connine and colleagues (Connine et al, 1991). They showed that a word with an initial phonetically ambiguous consonant midway between being heard as a voiced or voiceless consonant would be perceptually disambiguated by later context. For example, a sequence phonetically midway between "tent" and "dent", is reported as "tent" when followed by "….in the forest", and as "dent" when followed by "….in the fender". Bicknell et al (2015) report that the backward influence can extend over more than just the immediately following phrase (e.g., when the following context is either "….was noticed in the forest" vs. "….was noticed in the fender"). It is not clear from the methodologies used whether subjects believe they heard the critical word as disambiguated, or reasoned after the fact as to what the word must have been (For a discussion of these phenomena and related issues, see Bicknell et al (2015)).

The preceding cases involve the role of apparent "backwards" processing in which information that comes later in a sentence is used to specify or revise a prior analysis. A current line of experimental research by Brown, Dilley and Tanenhaus (2012) complements the study of conversational ellipses and the role of both forward and backward processing. In their study subjects think they "heard" a word that was acoustically ambiguous, or even marginally present at all, based on later acoustic input. Farmer, Brown and Tanenhaus (2013) apply Clark's (2013) model of hierarchically structured predictions to comprehension: the predictions guide the formation of representations of the world as new information becomes available.

"….Clark's framework predicts that expectations at higher levels of representation (e.g., syntactic expectations) should constrain interpretation at lower levels of representation (e.g., speech perception). According to this view, listeners develop fine-grained probabilistic expectations about how lexical alternatives are likely to be realized in context….that propagate from top to bottom through the levels of a hierarchically organized system representing progressively more fine-grained perceptual information….As the signal unfolds, then, the activation of a particular lexical candidate....[is the one] most congruent with the acoustic signal…." (Farmer, Brown, and Tanenhaus, 2013, p. 211)

This view of language comprehension emphasizes ongoing confirmation of hierarchically organized predictions, with error corrections when a given prediction is disconfirmed, shifting the interpretation of the prior material to an alternate hierarchical analysis. That is, material later in a sequence can revise the organization and interpretation of what came earlier, as a more subtle instance of the garden path phenomena explored by Ferreira et al. Brown et al. (2013 presented sentences with sequences like (3), and varied the length of the indefinite article, /a/ and the initial /s/ of the last word in the sequence. Using the 'visual world' paradigm, they report that when the article /a/ is shortened and the /s/ is lengthened, subjects look at plural target pictures ("raccoons") even after the /s/, indicating that the interpretation of the ambiguous noun in the sequence /a raccoon s….is/ determined on line by what follows it. That is, when the /s/ is lengthened, subjects first look at the picture with one raccoon; then as the lengthened /s/ is heard, they shift and look at the picture with several raccoons.

Ostensibly this reflects a reanalysis, in which the shortened /a/ is not treated as a separate word; it is attached as part of the final vowel of /saw/, or perhaps reanalyzed as a brief pause. This interpretation is strengthened by the complementary finding that when the /s/ is not lengthened, the shortened definite article is then perceived and interpreted as a word.

Brown et al's focus is on how their research shows that listeners are sensitive to variations in local speech rate, but for my purposes the phenomenon is an online demonstration of the use of late information in determining morphological analysis of earlier speech. (See also Brown et al (2012) for an example which arguably involves truly "hallucinating a definite article that was not present at all, based on extending the /s/). Importantly, Tanenhaus et al's view of how the comprehension of sentences proceeds is an example of a "top down" application of an interpretation, and perception in which an entire representation can be triggered by information at the end of the signal. This gives great weight to immediate access of contextual cues of a range of kinds, including actual syntactic hierarchical structure. (For more perspective on Tanenhaus's view on how representational levels interact during sentence comprehension, see Degen and Tanenhaus (2015)).

(3) ….saw uh raccoon Swimming

5. But in normal conversation, many words aren't there at all….

The preceding examples assume that all the words in the sentences are present to some degree. But in everyday speech, many acoustic details are slurred or even omitted. This can be demonstrated by showing that fragments several 'words' long are impossible to recognize in isolation, but pop into complete clarity (for native speakers) when heard as part of an entire sentence (Pollack and Pickett, 1964; Greenberg et al (1996); Arai (2009); Johnson, 2004; Warner *et al*., 2009; Tucker and Warner, 2010; For other

discussions of reduction in casual speech, see Ernestus (2000) Tucker and Warner, 2011; Ernestus and Warne, 2011; Dilley and Pitt, 2010; Gahl et al, 2012; and papers in the special issue of the journal of phonetics (2011), edited by Ernestus and Warner. Consider first an approximate transcription of an example from adults talking to each other in a normal conversation (this is an actual example provided by N. Warner, the reader can hear examples like it on her website:
http://www.u.arizona.edu/~nwarner/reduction_examples.html; also, readers interested in the examples discussed in this paper can email me for a powerpoint file with sound. tgb@:mail.arizona.edu)

(4) [tjutʌ̃m]

(Hint: this corresponds to four words). It is completely incomprehensible by itself, but when a latter portion of the longer sequence is included it is comprehensible:

(5) [tju tʌ̃m ɾɨ tʰak̚ tɨ̃ mi]

Everyone immediately hears this as:

(6) **Do you have time** to talk to me?

The striking significance of this is that consciously listeners think they simultaneously hear the fragment and assign it its three word analysis. But we know this cannot be true since the fragment in isolation is incomprehensible. This suggests that backward processing at a local acoustic level is a normal part of comprehension and building representations of conscious experience of language.

But this example was the beginning of a sentence, so perhaps it is a special case, where there is no preceding context. However in an experimental paradigm Van de Ven (2011) found that the following context can contribute importantly to recognition of material in the middle of a sentence. In fact, the following example from a natural conversation supports the view that in some cases, the following context alone is sufficient to clarify a reduced word, while the preceding context alone is not sufficient.

(7) [tʃɯ̃n:]

Try pronouncing this to yourself (hint: the production intent is 2 syllables). Now look at a longer sequence in which the example was embedded:

(8) [ɚ: ʌ: thɨzdɛ nʌit̚ (pause) ʌ̰ mn wɪɹ tʃɯ̃nɨ̃n:(ɨ) spa]

When listeners hear the surrounding material, the excerpt immediately pops into consciousness and what one 'hears' is:

(9) …err Tuesday night, when we were **chillin'** in the spa.

Recently we tested this further: it turns out that even with all the material preceding [tʃɯ̃n:] (as in "and err Tuesday night when we were….") almost no one perceives it correctly. But if only the following material ('in the spa') is heard along with the sequence, then [tʃɯ̃n:] is heard clearly as 'chillin'. First, such facts support the view that in everyday comprehension the minimal phonetic unit of comprehension is not the word, and that comprehension must be operating with parallel hypotheses at several interactive levels - syntactic and phonetic computations proceed in parallel with frequent cross checks at specific points. One can expect that where those cross checks occur will be the focus of ongoing research, now that we have tools that can chop running speech into a full range of possible units. An initial hypothesis is the *phase*, the unit of syntactic structure that has just enough content for semantic analysis (Chomsky, 1995, 2000, 2008). Phase theory is an active research area in linguistics, so the reader should be skeptical about details by the time this chapter is published, never mind a few years later. (See Boexck, 2006 for a lucid explication of the technical issues and Citko (2014) for a recent introduction.) So we can start with a particular hypothesis, as the latest idea on how different levels of a sentence are integrated in working units:

(10) *The unit over which local acoustic/phrasal/meaning integration occurs is the phase.*

However, we must note that "chillin" is involved in two prima facie phases: a) the preceding material which includes a WH, subject, and auxiliary, which embeds the verb in a complex structure with at least several levels of hierarchical organization; b) the following material, which embeds the verb in a more compact verbphrase only. The unique effectiveness of the following material leads to a hypothesis for further investigation, based on a single case, but one with some intuitive appeal:

(11) *The effectiveness of a phase in integrating distinct language levels is proportional to its structural simplicity.*

Further research will (I hope) verify or falsify these working hypotheses. A particular question is whether the role of the less complex phases is unique in the comprehension processes, or whether it reflects different degrees of reduction in the production processes. For example in (9) the failure of the preceding material to clarify the excerpt may be because as a NP-Verb phase it is actually less reduced in speech. So, it is now an interesting research question whether Phases are the 'true' units of comprehension that the many 'click' experiments attempted to define (Fodor & Bever, 1965; Garret *et al*., 1966; Bever *et al*., 1969), whether those effects depend on production processes, or whether the phase in fact is not the relevant factor that elicits segmentation effects. For example, there is new interest in how speakers maintain the predictability, (aka "information density") of their sentence output (e.g., Jaeger, 2006; Levy & Jaeger, 2007, Jaeger, 2010, Frank and Jaeger). This principle extends both to choice of phrases and words, and to use of contractions. For example, Frank and Jaeger show that local predictability can determine whether "you are" is contracted to "you're" in sentence production. Dell and Chang (2014) recently proposed a model that combines this approach with Macdonald's ideas that production patterns condition comprehension processes (Macdonald, 1999, 2013). Within a connectionist model of syntax production, they unify the processes of acquisition, production and comprehension based on serial predictability of words. The examples I have mentioned in this paper suggest that for such a model to be adequate, the unit of predictability is not only serial word-by-word, but ranges within a larger unit. It stands to reason that more complex phases (e.g., NP V) have more information and hence less internal predictability than simpler phases (e.g., VprepP). Thus, increased phonetic reduction in smaller phases (if true in general) could be due to structural or statistical factors in production. These alternatives open up the usual kind of research program in which a structural hypothesis (e.g., that the phase regulates speech production and phonetic reduction) competes with a statistical hypothesis (e.g., that units of mutual predictability regulate speech production and phonetic reduction). Specific experimental predictions are going to interact with each candidate theory of what phases are, so it is too rich an area to explore further here. But it does promise the possibility of an informative interaction between comprehension research, production research and using behavioral data to constrain theories of phases..

Implications for stages of comprehension and assigning syntax during processing. There is an intriguing interaction between the idea of analyzing serial sequences in whole chunks and Townsend's and my proposal about logical stages of alternating between associative and derivational processes during comprehension (Bever and Townsend, 2000; Townsend and Bever, 2001, chapters 5 and 8). We argued and reviewed evidence that comprehension processes necessarily integrate statistically valid patterns with computationally applied derivations, within an "analysis by synthesis" framework. On this model, pattern recognition templates can apply quickly to assign a likely meaning, to be complemented by derivational processes. This raised a question about when the derivational reconstruction of that input occurs: we answered this in the acronym for the model, LAST – "Late Assignment of Structure Theory" – making the point in contradistinction to other models which either assume that structure must be assigned before meaning, or that derivational structures are actually not assigned at all. In that work, most attention was given to the analysis of sentence level comprehension and syntactic structure assignment. The discussion in this paper gives some further organizational shape to the units within which pattern recognition and derivational processes can apply to assign meaning – our initial hypothesis for this is the phase, as described in (10). The demonstration of backward processes within such a unit supports the idea that comprehension proceeds in bursts that integrate learned patterns and composed structures.

The disconnect between unconscious processing and our conscious experience of normal conversational language calls into question the "immediacy assumption" – the theoretical preconception that a complete hierarchical layering of grammatical analyses is applied to language input *as we hear it* (Just and Carpenter (1980; Marslen-Wilson 1973, 1975). This assumption has been the bedrock of many distinct kinds of comprehension models (see Christiansen and Chater, 2016 for a review). The importance of backward processing of information I have reviewed shows that the assumption is false. I have focused on the importance of such processing for discriminating the speech signal. However, recent discussions have given a computational motivation for allowing indeterminate sequences to be held in immediate memory to be disambiguated or clarified by following input. On this model, there can be uncertainty immediately after each subsequence as to what it was: the subsequence is held in memory until the following material completes a larger pattern of analysis that embraces the critical subsequence (Levy et al, 2009; Kleinsschmidt & Jaeger; for general discussions se Kuperberg and Jaeger, 2015; Bicknell et al, ms). The critical point of consistency with the model in LAST is the notion that larger units immediately organize the local structure and ultimately the meaning of a lexical sequence. In the critical cases, an early indeterminacy is informed by its role in a larger unit of structure and meaning.

But this cannot be the whole story in the LAST framework. In our proposals, we noted that there must be a hierarchy of parallel levels during ongoing processing, each of which can involve integration of associative cues and structural assignments: this includes individual sounds, words, short phrases, phases, sentences and, arguably so-called "discourses" (see Townsend and Bever, 2001, Chapters 5 and 8; Bever and Poeppel, 2010). Integrating Clark's notion of parallel hierarchical processes with analysis-by-synthesis, we can think of these parallel computations as organized into a "computational fractal" in which the same alternation and integration of the two major kinds of information occur within each local linguistic unit (e.g., syllable, word, phrase, phase….): separate study of the processes at each level is a matter of "grain"- the size of each domain over which analysis by synthesis processing can occur.

This reinterpretation of our Analysis by Synthesis model moves towards a reconciliation between our view and the view that syntactic derivational structures are assigned serially from "left" to "right", as sentences are experienced. In this vein, Colin Philips has adduced arguments that such immediate structural assignment occurs, and also counter arguments to examples used by us to demonstrate the original analysis by synthesis proposals (for a review of his model and critique of ours, see e.g., Philips, 2013). In discussing our proposal, Philips also notes that an important issue is one of "grain". Our proposal here is that such processes occur in units of layered levels starting with individual sounds, overlapping with those of increasing size – that is, the processing is simultaneously multigrained. As it stands, this proposal offers a resolution of the theoretical conflicts, in principle, though much remains to be spelled out. And of course, it is important to review how Philips' positive research findings that support his model might also fit within the modified, "computational fractal" framework I am presenting here: but that will have to await a new thorough analysis.

6. Implications for notions of conscious experience

A related phenomenon is our conscious, but apparently false perception in many cases, that we understand the speech as we hear it serially. This has been long noted in phonology, but most of the effects are extremely local, and hence subject to solution by simply enlarging the scope of the initial input to a bigger chunk, e.g., the syllable, or word, as I mentioned. However, even in this case there is a puzzle: listeners 'think' consciously that they heard the individual sounds in words uttered in a citation form, in the order that they occurred. So, even at the most basic level of speech perception, our conscious experience of a series of stimuli, actually involves some "backward" processing.

The significance of this sort of phenomenon is magnified in the case of phrasal and sentence level processing. For example, in the cases of 'tyuv', and'chilln', where the critical (and incomprehensible) isolated sequence is *followed* by the crucial contextual material, we are not aware that we could not have analyzed the initial sequence until the later material was heard: rather we are convinced that we understood it as it was phonetically presented. This simple fact demonstrates that language comprehension may proceed in sequences of ‘psychological moments’ in which actual processing moves both forward and backwards, with some definition of phases specifying the domain of the interaction. This phenomenon has barely been touched in the language sciences, but is clearly fascinating and will have profound implications for consciousness theories, once it is better understood. Prima facie, it is an ultimate demonstration that even in language behavior (i.e. “externalization” of timeless linguistic structures) serial order may be less important than structure dependent organization.

There is a methodological dividend of the notion that there is a decoupling of the perceptual and comprehension processes and our consciousness of when they occurred. Throughout the literature on the post sentence location of clicks, when the reported location is not a phrase boundary, it systematically precedes the actual location (This started with Fodor and Bever, 1965, and has popped up several times, see also Townsend and Bever, (1991)). At first blush, this might be interpreted as a simple demonstration of the notion of “prior entry” (Titchener, 1908; Spence and Parise, 2009): an attended to stimulus is perceived earlier than others. It is possibly also related to demonstrations of “chronostasis” in which a more complex stimulus is slowed down relative to a simpler one. For example, Wundt reported a study in which a bell is perceived earlier than its actual location relative to a moving arrow across a series of numbers on a clock-face display. Wundt referred to the relative delay of the numbers as “positive time displacement” (Wundt, 1897, 1918). Correspondingly, in our studies, the subject’s task in locating the clicks is to locate the piece of the sentence and the click together, while attending to the entire sentence. To explain the preposition effect, we may refer to a Helmholtzian unconscious inference. Our conscious reconstruction of perceiving and understanding the speech stream as it was presented, leaves the click unanalyzed within the reconstruction of the speech. If it is the case that the click is perceived without the reconstruction processes, the unconscious inference is that it occurred earlier than it actually did. If one insists that this is merely an explanation of a well known “positive time displacement” or prior entry effect, at least it is an *explanation.*

The notion that conscious awareness of serial order can involve reconstruction is not novel. There is a distinguished line of research, stimulated by Husserl’s (1917,1990) considerations of the conscious perception of time, and most famously re-introduced by Fraisse (1967, 1974). However, most of the research in this vein involves relatively short intervals or rapid sequences of short and simple stimuli. For example, in demonstrations of metacontrast, a later stimulus will “absorb” an earlier one into an “exploding” or moving single object – indeed, this is a large part of how continuous motion is perceived in cinematic projections of at least 1 every tenth of a second. However, the language sequence cases described above involve much longer and more complex prospective and retrospective reconstructions. Thus, we have a potential demonstration that the “psychological moment” is itself determined by the perceptual units required: as they become more complex and hierarchical, the physical size of the “moment” can expand dramatically.

This concept returns us to the flagship issue of modularity in perceptual processing and representation, which Fodor famously explored. The corresponding puzzle for present and future research is how the distinct levels/modules of representation are actually integrated into the conscious experience of continuous integrated processing. That is, when I understand the sentence "a sentence is like a (miniature) opera" spoken conversationally, my conscious experience is that I hear and interpret the input as a coherent continuous object that unifies the acoustic input and the representational analysis; this occurs even though detailed examination of the sort I have reviewed here shows that the computational details belie this belief. In Fodor's formulation, the "central processor" is the mental cloaca where inputs and outputs to the different modules can meet: but, by definition, the central processor is relatively slow and woolgathering. So it remains to be spelled out how it could create the introspective belief that we understand sentences synchronously with their presentation. In Fodorian terminology, maybe it will turn out that consciousness itself is made up of the simultaneous output of a number of modules that interconnect with some degree of automaticity. As Fodor might say, stranger things have turned out to be true.

Thus, in this exploration, the study of language may become a theory-rich touchstone for yet another aspect of cognitive science – the nature of conscious experience.

7. The *Real* poverty of the stimulus

I began this discussion noting the significance of "the poverty of the stimulus" for all of cognitive science, as discussed by Fodor (1981).

Now consider the implications for the language-learning child of how sentences are acoustically mangled in normal conversation. There is evidence that child-directed 'motherese' is often clearer than normal conversations in many cases (Bernstein-Ratner, 1996; Bernstein-Ratner and Rooney, 2001), but not all (see Van de Weijer, 1991); it may use devices to clarify word boundaries (*e.g.*, Aslin *et al.*, 1996) and it may be that infants prefer motherese when they have a choice (*e.g.*, Fernald, 1985; Cooper *et al.*, 1997). In any case, it is likely that the vast majority of speech that children hear is between adults, or older children, and there are considerable cultural differences in whether motherese is used at all (Lieven, 1994). Furthermore, various studies have shown that the syntactic or phonetic quality of the child's input may bear little relation to the child's emerging language (C. Chomsky, 1986; McColgan, 2011). In any event, well-articulated motherese is not always dominant even in child-directed speech. Consider a transcribed example from a real motherese sentence. First, attempt to understand the fragment below (five words!), taken from an actual utterance by a mother to her child:

(11) [ɨ̃nʷ:ɹɨpə̃m]

Now see the whole utterance below. (If you are a phonetician) try sounding out the phonetic version alone to see if you can (suddenly) understand the whole utterance. In the acoustic version, adults cannot understand this sentence excerpt: but it immediately pops into perfect comprehension, with the conscious intuition that the entire utterance was reasonably clearly pronounced, which is immediately heard as in (15)

(14) [o gɹe(t) mamɪ mu ðoz mæɣəzɨ̃ns si jy k^{h}ɨ̃n: gɪɾɨ̃mɨ̃n^w:ɹɨpə̃m]

(15) Oh great, mummy put those magazines away so you can’t get them **and rip them**

It is amazing enough that adults can understand conversational speech like this. For a child the problem is doubly compounded, since its grammatical knowledge is incomplete, and it has not yet had time to build up complex language patterns. This simple fact vastly increases the poverty of the stimulus problem, since in many cases the child may not be able to even encode the utterance in enough detail to serve as a learning model.

There is an important implication of these analyses for how sophisticated the child’s comprehension system must be. Over many years, it has been argued that linguistic processes are *structure dependent* (Chomsky, 1980), rules are characteristically sensitive to hierarchical structure. This part of Universal Grammar has been shown to account for pathways to language in first language acquisition (e.g., Crain and Nakayama, 1987; and many others). Recent attempts have been made to show that serial learning models can converge on such sensitivity but such models fail to generalize realistically, omit structure dependence in fact (Perfors et al, 2006), or focus on simulating structure dependence (Reali and Christansen, 2005). (See Berwick et al, 2011, for general discussion). It has been shown that adults can learn serial rules but in so doing they utilize different brain areas than those characteristic of language (Musso et al, 2003; Moro, 2011). In the current “minimalist” treatments of language, hierarchical trees are constructed as sets, that is, without serial order constraints (Chomsky, 2007). On this view, the surface order in language is imposed by how it interfaces with our systems of input and output: but many actual computation of linguistic rules operate strictly on the hierarchical structures without reference to the serial structure of overt language sequences: thus, the comprehension system is building chunks of hierarchically organized structures which themselves may be internally order-free, corresponding to order free processing of the input.

Consider now, the implications of our idea that during language processing, there are “time free” processing zones that mediate between the serial input, structural analysis and simultaneous consciousness of the serial input and its meaning. Above, I suggest that the simplest available phase is the unit in which processing can occur both forward and backward. But this is to say in its strong form, that in certain defined domains, serial order is unconsciously suspended during sentence comprehension – allowing for structural dependencies to take precedence. In brief, within certain domains, even the externalization of language as serial may be ignored during behavior in favor of pure structure dependence.

A moment’s thought suggests that this must be so, as part of the solution to how the child manages to comprehend normal conversations and build up linguistic knowledge from them: s/he must be listening for phrasal categories that integrate and organize local word sequences. How else could s/he latch onto meanings and structural regularities so automatically and quickly? So, the argument that structure dependence appears spontaneously in children’s learning language structure applies perforce to early stages of language processing itself. (Cristophe et al 2008; For related discussion of models of how the language learning child might benefit from unsegmented input, see Pearl and Phillips 2016).

These considerations are consistent with an analysis by synthesis model of language acquisition, proposed in general terms in Bever, 1970, developed more specifically in Townsend and Bever, 2001, and elaborated in later writings (e.g., Bever, 2008; 2013). On this model, children alternate (logically) between accessing available structures/representational constraints and building generalizations over the language it experiences as represented by those categorical structures. The role of the generalizations is to provide form-meaning pairs for sentences that have not yet been assigned a full grammatical representation. These pairs can then be the input data for further elaboration of grammatical analysis, accessing the categorical structures. The categorical structures are in part innate - unique to language, in part innate as a part of general thought and perceptual processes. The categorical framework itself becomes more elaborate and uniquely adapted to language structure in particular. (See Bever, 2008, for further discussion and examples of this iterative process; Also Lidz and Gagliardi 2015, for a discussion of the interaction of generalizations and structure building during learning; see Bever, 2013, for a general discussion of this model of language acquisition as an instance of intrinsically motivated human problem solving).

The significant feature of this model is the dynamic integration of probabilistic and categorical information to yield both a repertoire of statistically valid generalizations and a constructed grammatical representation for all the sentences in the language and many of the semi-sentences. While the model has some general support from acquisition data, it is not sufficiently precise to be adequately testable in detail: in part this is because it is a framework for how associative and structural processes can interact, but allows for considerable differences between individuals and the data they experience.

Of course, this is not the first attempt to create a model of language behavior and acquisition that combines both associative and symbolic information. Indeed, the initial flowering of "psycholinguistics" under the leadership of Charles Osgood (Osgood and Sebeok, 1954; Osgood, 1968) was an explicit attempt to show that the then current model of mediated Stimulus-Response learning could account for the then current phrase structure model of language structure. (Alas, both models were inadequate for their respective goals, but were consonant with each other because the inadequacies corresponded well (See Bever, 1968, 1988, for discussions). In recent years, a class of explicit computational models has appeared that instantiates a dynamic integration of available categorical structures/processes and Bayesian inference algorithms. These models ostensibly connect to Fodor's notion of the Language of Thought (LOT), the set of symbols and processes that manipulate symbols. The recent models add a Bayesian statistical component to LOT, and recast it as the "Probabilistic Language of Thought" (pLoT). Researchers in this vein show that many graded kinds of category knowledge can be accounted for as well as apparent category and concept formation. (See Perfors et al, 2006; Goodman and Lassiter (2014); Piantadosi and Jacobs, 2016 for representative discussions among the many articles now appearing on pLoT). It remains to be seen if such models can actually learn or even render grammatical representations, including processes that involve structure dependent constraints. At the moment these models do not generally address such problems.

This is not to say that no attention is given to how statistically non-categorical input can result in arriving at grammars appropriate to the child's native language. A number of models have also used Bayesian and other statistical techniques of how variable input data may discriminate between candidate grammars. This includes many different target architectures, but all in the general method of using statistically variable input to reinforce or distill out candidate rules or grammars. (e.g., See Yang, 2002, 2004; Yang and Roeper 2011; Pearl and Goldwater 2016; Lidz and Gagliardi, 2015). The critical feature that seems to discriminate these approaches from the emerging pLoT variants of Fodor's "Language of Thought", is that these approaches presuppose the availability of candidate grammars or rules, both in the child and as the ultimate goal of language learning.

Implications for old and new research.

A cautionary note on the issue of how children and adults deal with normal conversational speech: *sometimes* our spoken utterance may be clear enough, with adequate serial cues for a diligent listener to develop immediate representations of what s/he is hearing. This *may* be especially true of instances of so called child-directed "motherese". But what is important in our examples is that this is not always the case, and may not even be true of the majority of cases. If indeed, most of our comprehension has to deal with compressed and cue-poor input, this also calls into question the generalizability of the many studies of carefully pronounced "laboratory speech" that comprise the overwhelming majority of experimental studies, never mind the use of complete word-by-word visual presentation.

The reader will note that I have extrapolated very far ahead of a very small number of facts, but I hope in ways that are amenable to empirical investigation. For example, one can use the backwards reconstruction phenomenon as a tool to study what units are the relevant bridges between serial input and structural output. Here is a(n in principle) simple way to do this. Take conversational corpora and analyze the transcripts (which presumably already have interpreted the conversations into complete words, phrases and sentences); pick out candidate phases according to a theory of what phases are relevant (e.g., as postulated in (10)); test gated increments of each candidate from its beginning for recognition of the input by subjects (that is, start with an initial fragment, then successively longer ones to see when the initial fragment becomes (retrospectively) clearly interpretable; do the corresponding testing starting from the final part of such fragments. The same kind of procedure can be applied to child-directed speech to examine empirically the claim that a great deal of it is also heavily encoded and dependent on both forward and backward processing. No doubt, these are big projects, but the payoff could be even bigger in leading to a theoretical understanding of how serially presented units build up hierarchical structures and meaning in comprehension and language learning, and to information about normal speaking with many practical applications.

Such research programs can be viewed as the latest step in making good on the implications of the original discoveries by Fodor and his colleagues that in ongoing speech comprehension, sentences are automatically segmented into natural units.

References

Abrams, K., & Bever, T. G. (1969). Syntactic structure modifies attention during speech perception and recognition. *Quarterly Journal of Experimental Psychology*, *21*, 280-290.

Arai, T. (1999). “A case study of spontaneous speech in Japanese,” in Proceedings of the International Congress of Phonetic Sciences, San Francisco, 1, 615–618.

Arai, T. & Warner, N. (1999). Word Level Timing in Spontaneous Japanese Speech. In *Proc. Int’l Cong. on Phonetic Sciences*.

Aslin, R. N., Woodward, J. Z., LaMendola, N. P., & Bever, T. G. (1996). Models of word segmentation in fluent maternal speech to infants. In J. L. Morgan, K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp.117-134). Hillsdale, NJ: Erlbaum.

Bernstein-Ratner, N. (1996). From signal to syntax: But what is the nature of the signal. In J. L. Morgan, & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 135-150). Hillsdale, NJ: Erlbaum.

Bernstein- Ratner, N., & Rooney, B. (2001). How accessible is the lexicon in motherese? In J. Weissenborn., & B. Höhle (Eds.), *Approaches to bootstrapping: Phonological, lexical, syntactic and neurophysiological aspects of early language acquisition Vol. I.* (pp.71-78). Philadelphia: Benjamins Publishing Company.

Berwick, R. C., Pietroski, P., Yankama, B., & Chomsky, N. (2011). Poverty of the stimulus revisited. *Cognitive Science*, *35*(7), 1207-1242.

Bever, T.G. (1968). Associations to stimulus-response theories of language. in T.R. Dixon & D.L. Horton (Eds.), *Verbal behavior and general behavior theory* (478-494). Prentice-Hall, Inc.

Bever, T.G., (1970, 2013) The Cognitive Basis for Linguistic Structures. Reprinted in Laka, I. ,Sanz, M. and Tanenhaus, M. (Eds), *Cognition and Language: enduring themes*. Oxford University Press. pp.1-70.

Bever, T. G. (1973). Serial position and response biases do not account for the effect of syntactic structure on the location of brief noises during sentences. *Journal of Psycholinguistic Research. 2*, 287-288.

Bever, T.G. (1988). The psychological reality of grammar: a student's eye view of cognitive science. In W. Hirst (Ed.), *The making of cognitive science* Cambridge

University Press.

Bever, T.G., (2008) The canonical form constraint: language acquisition via a general theory of learning. In Guo et al, *Cross-linguistic approaches to the psychology of language.* (Oxford University Press) pp.475-492

Bever, T. G., (2009) Minimalist Behaviorism: the role of the individual in explaining language universals. In Christiansen, M. Collins & Edelman, S. (Eds.) *Language Universals* (Oxford University Press). Pp. 270-298

Bever, T.G (2013, 2015), The biolinguistics of language universals: The next years. In Montserrat Sanz, Itziar Laka, Michael K. Tanenhaus (Eds,) *Language Down the Garden Path: The Cognitive and Biological Basis of Linguistic Structures*, Oxford University Press. Pp. 235-405.

Bever, T. G., Lackner, J. R., & Stolz, W. (1969). Transitional probability is not a general mechanism for the segmentation of speech. *Journal of Experimental Psychology*, *79*(3), 387-394.

Bever, T. G., Lackner, J. R., & Kirk, R. (1969). The underlying structures of sentences are the primary units of immediate speech processing. *Perception & Psychophysics*, *5*(4), 225-234.

Bever, T. G., & Poeppel, D. (2010). Analysis by Synthesis: A (Re-)Emerging Program of Research for Language and Vision. *Biolinguistics*, *43*(2), 174–200.

Bicknell, K., Jaeger, T. F., & Tanenhaus, M. K. (In press). Now or . . . later: Perceptual data is not immediately forgotten during language processing. *Behavioral and Brain Sciences*.

Bicknell, K., Tanenhaus, M. and Jaeger, T., Rationally updating uncertainly about previous words. Ms.

Boeckx, C. (2006). *Linguistic Minimalism: Origins, Concepts, Methods, and Aims.* New York: Oxford University Press.

Brown, M., Dilley L. C., & Tanenhaus, M. K. (2012). Real-time expectations based on context speech rate can cause words to appear or disappear. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (pp. 1374-1379). Austin, TX: Cognitive Science Society.

Citko, B. (2014). *Phase theory: An introduction.* Cambridge, UK: Cambridge University Press.

Chapin, P. G., Smith, T. S., & Abrahamson, A. A. (1972). Two factors in perceptual segmentation of speech. *Journal of Verbal Learning and Verbal Behavior*, *11*(2), 164-173.

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

Chomsky, C. (1986). Analytic study of the tadoma method: Language abilities of three deaf/blind children. *Journal of Speech, Language and Hearing Research*, *29*(3), 332-347.

Chomsky, N. (1959). A review of BF Skinner's verbal behavior. *Language*, *35*(1), 26-59.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Oxford, England: M.I.T. Press.

Chomsky, N. (1975). *Reflections on language*. NewYork, NY: Pantheon Books.

Chomsky, N. (1995). *The minimalist program.* Cambridge, MA: Massachusetts Institute of Technology.

Chomsky, N. (2000). *New horizons in the study of language and mind.* Cambridge, UK: Cambridge University Press.

Chomsky, N. (2008). On Phases. In R. Freidin, C, Peregrín Otero, & M. L. Zubizarreta (Eds.), *Foundational Issues in Linguistic Theory. Essays in Honor of Jean-Roger Vergnaud* (p. 133-166). Cambridge, MA: MIT Press.

Christianson, K., Hollingworth, A., Halliwell, J. F., & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology*, *42*, 368-407.

Christianson, K., Williams, C. C., Zacks, R. T., & Ferreira, F. (2006). Younger and Older Adults' 'Good-Enough' Interpretations of Garden-Path Sentences. *Discourse Processes*, *42*(2), 205-238.

Christophe, A., Millotte, S., Bernal, S., & Lidz, J. (2008). Bootstrapping lexical and syntactic acquisition. *Language and speech*, *51*(1-2), 61-75)

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(03), 181-204.

Cooper, R. P., Abraham, J., Berman, S., & Staska, M. (1997). The development of infants' preference for motherese. *Infant Behavior and Development*, *20*(4), 477-488.

Connine, C.M., Blasko, D.G., & Hall, M. (1991). Effects of subsequent sentence context in auditory word recognition: Temporal and linguistic constraints. Journal of Memory and Language, 30, 234-250.

Crain, S., & Nakayama, M. (1987). Structure dependence in grammar formation. *Language*, *63*(3), 522-543.

Dalrymple-Alford, E. C. (1976). Response bias and judgments of the location of clicks in sentences. *Perception & Psychophysics*, *19*(4), 303-308.

Dell GS, Chang F. (2014) The P-chain: relating sentence production and its disorders to comprehension and acquisition. Phil. Trans. R. Soc. B 369

Dilley, L., & Pitt, M. A. (2010). Altering context speech rate can cause words to appear and disappear. *Psychological Science*, *21*(11), 1664-1670.

Ding, N, Melloni, L., Zhang, H. Tian, X & David Poeppel (2016) Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience.* 19(1), pp. 158-164

Ernestus, M. (2000). *Voice Assimilation and Segment Reduction in Casual Dutch: A Corpus-Based Study of the Phonology-Phonetics Interface* [dissertation]. Utrecht: LOT.

Ernestus, M., and Warner, N., eds. (2011). Special issue of *Journal of Phonetics* on speech reduction. **39(3)**. Pp. 1-14

Farmer, T. A., Brown, M., & Tanenhaus, M. K. (2013). Prediction, explanation, and the role of generative models. *Behavioral and Brain Sciences*, *36*, 211-212.

Farmer, T. A., Yan, S., Bicknell, K., & Tanenhaus, M. K. (2015). Form-to-expectation matching effects on first-pass eye movement measures during reading. *Journal of Experimental Psychology: Human Perception and Performance 41*, 958–976.

Ferreira, F., Bailey, K. D., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions In Psychological Science*, *11*(1), 11-15. doi:10.1111/1467-8721.00158.

Ferreira, F., & Henderson, M. (1991). Recovery from misanalyses of garden-path sentences. *Journal of Memory and Language*, *31*, 725-745. doi:10.1016/0749-596X(91)90034-H.

Ferreira, F., & Patson, N.D. (2007). The 'good enough'approach to language comprehension. *Language and Linguistics Compass*, *1*(1‐2), 71-83.

Ferreira, F., Engelhardt, P. E., & Jones, M. W. (2009). Good enough language processing: A satisficing approach. In N. Taatgen, H. Rijn, J. Nerbonne, & L. Schomaker (Eds.), *Proceedings of the 31st Annual conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society**.**

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development*, *8*(2), 181-195.

Fodor, J. A. (1981). *Representations: Philosophical Essays on the Foundations of Cognitive Science*. Cambridge, Mass.: The MIT Press.

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

Fodor, J. A., Bever., T. G., & Garrett, M. F. (1974). The psychology of language: An introduction to psycholinguistics and generative grammar. New York: McGraw-Hill.

Fraisse, P. (1967). *Psychologie du temps.* Paris: Presses universitaires de France.

Fraisse, P. (1974). *Psychologie du rhythme*, Paris: Presses universitaires de France.

Frank A, Jaeger TF. Speaking rationally: Uniform information density as an optimal strategy for language production. The 30th annual meeting of the Cognitive Science Society (CogSci08); Washington, DC. 2008. pp. 939–944.

Gahl, S., Yao, Y., & Johnson, K. (2012). Why reduce? Phonological neighborhood density and phonetic reduction in spontaneous speech. *Journal Of Memory And Language*, *66*(4), 789-806.

Garrett, M. F. (1964). *Structure and sequence in judgments of auditory events* (Unpublished doctoral dissertation). University of Illinois, Chicago.

Garrett, M. F., Bever, T. G., & Fodor, J. A. (1966). The active use of grammar in speech perception. *Perception & Psychophysics*, *1*, 30-32.

Goodman and Lassiter (2014), Probabilistic Semantics and Pragmatics; uncertainty in language and thought. In Lappin (S. and Fox, C. *Handbook of contemporary semantics.* Wiley-Blackwell

Greenberg, S., Hollenback, J., & D. Ellis (1996); Insights into spoken language gleaned from phonetic transcription of the Switchboard corpus. In *Proc of ICSLP*, pages 24–27, Philadelphia.

Greenberg, S. (1999) Speaking in shorthand – a syllable-centric perspective for understanding pronunciation variation. *Speech Communication*, 29:159–176.

Henderson, J. M., & Ferreira, F. (2004). *The interface of language, vision, and action: Eye movements and the visual world*. New York, NY, US: Psychology Press.

Husserl, E. (1917). On the phenomenology of the consciousness of internal time (1893-1917) (Vol. 4) (J. B. Brough, Trans.). Dordrecht: Kluwer (Original work published in 1991).

Johnson, N. F. (1970). The role of chunking and organization in the process of recall. *Psychology of Learning and Motivation*, *4*, 171-247.

Johnson, K. (2004) Massive reduction in conversational American English. In K. Yoneyama & K. Maekawa (eds.) *Spontaneous Speech: Data and Analysis. Proceedings of the 1st Session of the 10th International Symposium.* Tokyo, Japan: The National International Institute for Japanese Language. pp. 29-54.

Jaeger, T. F. (2006). *Redundancy and syntactic reduction in spontaneous speech*. Unpublished doctoral dissertation, Stanford University.

Jaeger TF. (2010_. Redundancy and reduction: speakers manage syntactic information density. Cogn. Psychol. 61, 23 – 62

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: from eye fixations tocomprehension. Psychological Review, 87(4), 329–354.

Kleinschmidt, D. F., & Jaeger, T. F. (2015). Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. Psychological Review, 122(2), 148–203.

Ladefoged, P., & Broadbent, D., Perception of Sequence in Auditory Events. Quart. J. exp. Psychol., 1960, 12, 162-170.

Levy, R., Bicknell, K., Slattery, T., & Rayner, K. (2009). Eye movement evidence that readers maintain and act on uncertainty about past linguistic input. Proceedings of the National Academy of Sciences, 106(50), 21086-21090

Levy, R., & Jaeger, T. F. (2007). Speakers optimize informa- tion density through syntactic reduction. In B. Scho ̈lkopf, J. Platt, & T. Hoffman (Eds.), (p. 849-856). Cambridge, MA: MIT Press.

Lidz, J. and Gagliardi, A. (2015) How Nature Meets Nurture: Universal Grammar and Statistical Learning Annual Review of Linguistics Vol. 1: pp. 333-353

Lieven, E. V. M. (1994). Crosslinguistic and crosscultural aspects of language addressed to children. In C. Gallaway & B. J. Richards (*Eds.*), *Input and Interaction in Language Acquisition* (pp. 56-73). Cambridge: CUP.

MacDonald MC. (1999). Distributional information in language comprehension, production, and acquisition: three puzzles and a moral. In *The emergence of language* (ed. B MacWhinney), pp. 177 – 196. Hillsdale, NJ: Lawrence Erlbaum Associates.

MacDonald MC. (2013). How language production shapes language form and comprehension. *Front. Psychol*. 4, 226..

Marslen-Wilson, W. (1973). Linguistic structure and speech shadowing at very short latencies. *Nature,* 244, 522–523.

Marslen-Wilson, W. (1975). Sentence perception as an interactive parallel process. *Science*, 189, 226–228.

Marslen-Wilson, W., & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal Of Experimental Psychology: Human Perception And Performance*, *15*(3), 576-585.

McColgan, K. (2011). *The Relationship between Maternal Language Outcomes* (Unpublished masters thesis). University of Maryland, Maryland.

Miller, G. A. (1962). Some psychological studies of grammar. *American Psychologist*, *17*(11), 748-762.

Moro, A. (2011). A closer look at the turtle's eyes. *Proceedings of the National Academy of Sciences*, *108*(6), 2177-2178.

Musso, M., Moro, A., Glauche, V., Rijntjes, M., Reichenbach, J., Büchel, C., & Weiller, C. (2003). Broca's area and the language instinct. *Nature Neuroscience*, *6*(7), 774-781.

Osgood, C. E. (1968). Toward a wedding of insufficiencies. In T. R. Dixon, & D. L. Horton (Eds.), *Verbal Learning and General Behavior Theory* (pp. 495-519). Englewood Cliffs: NJ: Prentice-Hall.

Osgood, Charles, and Thomas Sebeok. 1954. Psycholinguistics: A Survey of Theory and Research Problems. *J. Abn. and Social Psych*. 49 (4, part 2) : 1-203.

Pearl, L. & Goldwater, S. 2016. Statistical Learning, Inductive Bias, and Bayesian Inference in Language Acquisition. In J. Lidz, W. Snyder, & C. Pater (eds), *The Oxford Handbook of Developmental Linguistics*, 664-695.

Pearl, L. & Phillips, L. (in press). Evaluating language acquisition models: A utility-based look at Bayesian segmentation. In A. Villavicencio & T. Poibeau (eds), *Language, Cognition and Computational Models*, Cambridge University Press.

Piantadosi, S. and Jacobs, R. (2016) Four problems solved by the probabilistic language of thought. *Current directions in psychological science.* Vol. 25(1) 54-59

Perfors, A., Tenenbaum, J., & Regier, T. (2006). Poverty of the Stimulus? A rational approach. In R Sun and N Miyake (Ed.) *Proceedings of the 28th Annual Conference of the Cognitive Science Society* (pp. 663-668). Mahwah, NJ: Lawrence Erlbaum Associates.

Poeppel, D., Idsardi, W. and von Wassenhjove, V. (2007) Speech perception at the interface of neurobiology and linguistics. Phil. Trans. R. Soc. B doi:10.1098/rstb.

Pollack, I., & Pickett, J. M. (1964). Intelligibility of excerpts from fluent speech: Auditory vs. structural context. *Journal of Verbal Learning and Verbal Behavior*, *3*, 79-84.

Reali, F., & Christiansen, M. H. (2005). Uncovering the richness of the stimulus: Structural dependence and indirect statistical evidence. *Cognitive Science*, *29*(6), 1007-1028.

Rumelhart, D. E., McClelland, J. L., & PDP Research Group (1986). Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volumes I & II, Cambridge, MA: MIT Press.

Slattery, T. J., Sturt, P., Christianson, K., Yoshida, M., & Ferreira, F. (2013). Lingering misinterpretations of garden path sentences arise from competing syntactic representations. *Journal of Memory and Language*, *69*(2), 104-120.

Spence, C., & Parise, C. (2009). Prior-entry: A review. *Consciousness and Cognition*, *19*(1), 364-379.

Swinney, D. (1979), Lexical access during sentence comprehension:(Re) consideration of context effects. *Journal Of Verbal Learning And Verbal Behavior* **18**, 645-659 (1979)

Titchener, E. B. (1908). Lectures on the elementary psychology of feeling and attention. New York: Macmillan.

Townsend, D., & Bever, T. G. (1991). The use of higher-level constraints in monitoring for a change in speaker demonstrates functionally distinct levels of representation in discourse comprehension. *Language and Cognitive Processes*, *6*(1), 49-77.

Tucker, B. V., & Warner, N. (2010). What it means to be phonetic or phonological: The case of romanian devoiced nasal. *Phonology*, *27*(2), 1-36.

Van de Ven, M.A.M. (2011). he role of acoustic detail and context in the comprehension of reduced pronunciation variants. PhD Thesis, Radboud Universiteit Nijmegen,

Van de Weijer, J. (1998). Language Input for Word Discovery. Ph.D. thesis. Max Planck Series in Psycholinguistics, 9.

Warner, N., Fountain, A., & Tucker, B. V. (2009). Cues to perception of reduced flaps. *Journal of the Acoustical Society of America*, *125*(5), 3317-3327.

Wundt, W. (1918) *Grundriss der Psychologie* Alfred Kroener Verlag, Leipfzig.

Yang, C. (2002) Knowledge and Learning in Natural Language, Oxford University Press

Yang, C.(2004) Universal Grammar, statistics or both? *TRENDS in Cognitive Sciences* Vol.8 No.10 October 2004 pp.451-456.

Yang, C. and Roeper, T. Minimalism and Language Acquisition (2011). in Boeckx, Ed. *The Oxford Handbook of Linguistic Minimalism*

Footnotes

1. It should be obvious to the reader how much this paper owes to Jerry Fodor. Along with Merrill Garrett, we pioneered click mislocation as a method to demonstrate the active on-line use of syntax during comprehension: this is the foundation of many subsequent explorations of the initial compositional strategies of comprehension. More personally, my many conversations with Jerry, co-teaching a course in the early 1960s, and co-authoring our 1974 book ("The Psychology

of Language"), gave me wide-ranging instructions in how to think about the general problems of cognitive science. We did discuss the poverty of the stimulus, both in relation to adult comprehension and language acquisition. But we did not discuss consciousness at all, to my recollection: it was viewed at the time as a slightly embarrassing romantic problem not a scientific one. And Jerry was never convinced about the value of phenomenology (e.g., Husserl, Heidiger): he is reputed to have quipped about Heidigerian Dasein: "I don't know what it is, but I bet my cat and Chicago have it". On consciousness more generally Jerry noted in his 2007 review of Strawson's edited book, "[it] is all the rage just now….What everybody worries about most [is] what philosophers have come to call 'the hard problem'. The hard problem is this: it is widely supposed that the world is made entirely of mere matter, but how could mere matter be conscious? How, in particular, could a couple of pounds of grey tissue have experiences?" In considering this question, I (TgB) follow the general approach in "biolinguistics" to an understanding of the biology and genetics of language: to discover what makes consciousness possible, we first have to determine what consciousness ***is***, how it is acquired as a habit from the alleged "blooming buzzing confusion" of infancy, how it is represented, how it works. This paper is not a solution to all that, but a pointer to a problem that I hope will attract the interest of today's graduate students. Without them, our science will be lost to the world.

I am indebted to Roberto Almeida and Lila Gleitman, editors of the volume where this appeared, for many helpful early criticisms and comments; also to Mahmoud Azaz for conceptual advice, to David Poeppel for reminding me about the broad evidcnce for what I coin the "computational fractal" in language processing, and especially to Michael Tanenhaus for deeply considered advice. Caitlyn Antal and Valerie Kula were invaluable bibliographic assistants. Other helpful comments are due to; Stuart Hameroff, Al Bergesen, Felice Bedford, Virginia Valian, Rudi Troike, Louann Gerken, Billl Idsardi, Gary Dell, Florian Jaeger, Mark Pitts, Maryellen MacDonald, Lisa Pearl, Massimo Piattelli-Palmarini and Noam Chomsky. Most of all, I am indebted to my colleague Natasha Warner, who not only developed a methodology of collecting natural conversations, but also made available her materials, her bibliographic advice, her perspective, and all the phonetic transcriptions in this paper.