The Interplay of Meaning, Sound, and Syntax in Sentence Production

Gabriella Vigliocco
University College London

Robert J. Hartsuiker
University of Edinburgh

A discussion of modularity in language production processes, with special emphasis on processes for retrieving words and building syntactic structures for a to-be-uttered sentence, is presented. The authors' 1st goal was to assess the extent to which information processing is encapsulated between different processing stages. In particular, they assessed whether the input from one processing stage to the next is minimal and whether the flow of information in the system is strictly unidirectional. On the basis of the reviewed evidence, they conclude that both assumptions have to be revised. Their 2nd goal was to propose an alternative framework that does not assume strict encapsulation but that maintains multiple levels of integration for production.

During the past 20 years a "great divide" (as Boland & Cutler, 1996, p. 309, labeled it) has characterized the psycholinguistic world: Can the processes engaged during comprehension and production of language be conceived of as modular or not? Whereas a number of influential theories that do not assume a modular system have been put forward in the comprehension domain (e.g., MacDonald, Pearlmutter, & Seidenberg, 1994; Tabor & Tanenhaus, 1999; Tanenhaus & Trueswell, 1995) motivated on evidence indicating that comprehenders use all the available information as the speech signal unfolds (including phonological, metrical, syntactic, and crucially semantic information), the production system is often considered to be an almost paradigmatic example of a modular system (Garrett, 2000; Levelt, 1989) in which processes at one level (e.g., syntactic encoding) are encapsulated from the processes occurring at other levels (e.g., phonological encoding). Note that this is somewhat peculiar because the notion of modularity did not apply to subsystems within language in Fodor's view (1983, 1985) and also because output systems were not included among the modules; only the input systems were. In fact, with respect to language production, Fodor (1983) wrote as follows:

> We use language to communicate our views on how the world is. But this use of language is possible only if the mechanisms that mediate the production of speech have access to what we see (or hear, or remember, or think) that the world is like. Since, by assumption, such mechanisms effect an interface among vertical faculties, they cannot themselves be domain specific. More precisely, they must be *less* domain specific than the vertical faculties are. (p. 102)

In this article we assess the modularity claim as applied to sentence production by closely examining a number of interfaces between processing levels (which we henceforth call *joints*) in the production system. After a general introduction in which we discuss some basic properties of sentence production, we provide an overview of standard, modular models of sentence production (as proposed by Garrett, 1976, 1982, 2000; Levelt, 1989, Levelt, Roelofs, & Meyer, 1999). We then present evidence that challenges predictions derived from these models, and we sketch in broad strokes an alternative view that accommodates these data, a view that combines assumptions from the theories developed by Bates and MacWhinney (1982, 1989), Dell (1986), Kempen and colleagues (Kempen & Vosse, 1989; Kempen, 1999; Vosse & Kempen, 2000), MacKay (1992), Trueswell, Tanenhaus, and Garnsey (1994), and Tanenhaus and Trueswell (1995). We conclude by discussing some important similarities and differences between sentence production and sentence comprehension.

MINIMALISM, MAXIMALISM, AND LEVELS OF INTEGRATION

Two fundamental properties of language production that theories ought to account for are its *accuracy* and its *efficiency*. Speakers produce on average 2–3 words per second. These words are retrieved from a lexicon of approximately 30,000 (productively used) words. This is no small feat: Producing connected speech entails much more than retrieving words from memory; it also entails combining stored information and constructing, on the fly, syntactic relations among the words (e.g., agreement between the subject of the sentence and the verb). Furthermore, speech production entails combining information from memory to obtain phonological words (e.g., the simple phrase "demand it" involves two lexical units *demand* and *it*, but in running speech these two are combined into a single phonological word *demandit*) and coarticulation (i.e., phonemes are pronounced differently in differ-

ent phonetic environments; /p/ in *spin* is different from /p/ in *pin*) to mention just a subset of the processes involved in this highly complex activity. Given the complexity of all these encoding processes, it is remarkable that people produce speech at such a fast rate. It is also remarkable that we are so very accurate in our production. Bock (1991) reported that slips of the tongue occur approximately once every 1,000 words, as estimated from the London–Lund corpus (Garnham, Shillcock, Brown, Mill, & Cutler, 1981). This level of efficiency and accuracy may arise in two ways. These properties may be driving forces in the development of the architecture, or alternatively, they may be a by-product of a given architecture. Whichever of these alternatives is the case, a theory of sentence production should account for the accuracy and efficiency of the system.

In general terms, we can consider theories in two broad classes with respect to issues of accuracy and efficiency (Vigliocco & Franck, 1999, 2001). One class of theories embraces modularity (Fodor, 1983) and, in particular, the assumption of *information encapsulation,* defined as "not having access to facts that other systems know about" (Fodor, 1983, p. 73), as applied to different levels (syntactic, phonological) of processing in language production (e.g., Garrett, 1975, 1980, 2000; Levelt, 1989; Levelt et al., 1999). We label this view as *minimalist,* because the information flow among levels is characterized by the minimal necessary information (shallow output of a module). A further claim of this view is that information processing is unidirectional. In other words, there is no interaction among levels. In this framework, accuracy is achieved by insulating processes at one level (e.g., syntactic encoding) from nonnecessary information at other levels (e.g., conceptual and phonological) that could potentially interfere. Efficiency also derives from these fundamental properties of modular systems: Because modules are narrowly focused, they can be fast in their functioning (Fodor, 1983, p. 61).

Another class of theories lies on the other side of the great divide (Boland & Cutler, 1996). These theories embrace interactivity among different information types (e.g., Dell, 1986; Harley, 1993; Stemberger, 1985). Among these theories we also have functionalist models, such as the competition model proposed by Bates, MacWhinney, and colleagues (e.g., Bates & MacWhinney, 1989; Bates, McNew, MacWhinney, Devescovi, & Smith, 1982), that consider language use as tailored by communication drives on the one hand and by the structural properties of the specific languages used on the other. We label this general view as *maximalist,* because at each level converging information from other levels, including later levels, could exert an effect. In a maximalist framework, accuracy is achieved by using the converging information available to the system. For example, if a syntactic source of information is lost, additional nonsyntactic information (correlated with the syntactic information) could be helpful in compensating for this loss. Efficiency in this view may result as a consequence of maximal input. For example, a type of maximal input such as *cascading of activation* (i.e., spreading of activation to a given level before a selection is achieved at the previous level) allows for faster encoding by virtue of preactivation (Dell, Burger, & Svec, 1997). Bidirectional flow of information may also subserve efficiency (in the sense of fluency). Dell, Burger, and Svec (1997), for example, suggested that at the lexical level feedback may help to preserve fluency by providing a dynamic mechanism by which processes at one level are informed about the retrievability of corresponding information at a subsequent level. This would support fluency because it would reduce the probability of lexical retrieval blockages, such as tip-of-the-tongue (TOT) states.

In summary, with respect to accuracy, the contrast between minimalist and maximalist theories is set in terms of whether converging information from different levels interferes or instead may protect from loss and underspecification. Therefore, the two views differ with respect to which potential source of errors the system's architecture shields us from: interference in the minimalist view and information loss and underspecification in the maximalist view. With respect to efficiency, the minimalist theory's assumptions of minimal input and unidirectionality of processing guarantee that the operations of each module are fast. In the maximalist theories, efficiency is ensured by assuming cascading of activation (a specific form of maximal input), which allows for preactivation of upcoming units, and by feedback.

Newell (1973) warned us that "you can't play twenty questions with nature and win" (p. 283). Our way of interpreting this warning is that one should not pose yes–no questions to nature without embedding them in a theoretical framework. This also holds for our questions with respect to the minimalist and maximalist accounts of accuracy and efficiency. In this article we use as our framework a theory of sentence production that was first proposed and developed by Merrill Garrett (1975, 1980, 1982) and whose basic architectural assumptions have been inherited by many more recent theories (e.g., Bock & Levelt, 1994; Levelt, 1989). We review the evidence with respect to this (minimalist) theory's architecture. Note that with respect to lexical retrieval processes, our evaluation of minimalist assumptions goes beyond Garrett's general proposal to assess the computational theory of lexical access proposed by Levelt et al. (1999). Below, then, we provide the architectural constraints within which our questions to nature are asked.

There is now a plethora of evidence for distinct processing levels that integrate stored linguistic information into sentential frames (which we review in the next section). In particular, three broad levels of representation (and corresponding processes) are usually postulated in going from "mind to mouth" (Bock, 1995): the message level (Garrett, 1982), the domain of the conceptualizer (Levelt, 1989); the sentence level (Garrett, 1982), the domain of the formulator (Levelt, 1989); and the articulatory level, the domain of the articulator. The sentence level of processing is further divided into different sublevels: the functional, positional, and phonetic levels (Garrett, 1984).

Evidence compatible with such a "levels-of-integration" approach is reviewed below in the overview of theories of sentence production that embody minimalism. There, we also review the available evidence that suggests minimal input and unidirectionality among processing components. Problematic data for a minimalist levels-of-integration view is presented in the following section. We then present an alternative framework that maintains the levels of integration of the minimalist view but allows for a greater degree of cross talk among information types (maximalist view). We conclude by briefly considering the modularity thesis in the related domain of language comprehension.

A MINIMALIST LEVELS-OF-INTEGRATION APPROACH TO SENTENCE PRODUCTION

Figure 1 provides a sketch of the processes and representations involved in sentence production according to Garrett (1984). Below we describe and discuss the different components of this system.

In our presentation, we first review evidence for assuming different levels of integration during sentence production. Subsequently, we discuss evidence compatible with minimalism at the different joints in the system as has been argued by Garrett (e.g., 1976; 2000) and Levelt et al. (1999). We separately address the two key aspects of minimal flow of information from a subsystem to another: (a) *minimal input,* that is, a process at a level *n* passes only minimal, necessary information to a level *n* + 1, and (b) *unidirectional flow of information,* that is, a process at a level *n* + 1 does not feed back information to a level *n*. With respect to this second assumption we contrast the minimalist claim of unidirectional flow to a maximalist view that allows for feedback but that is only locally interactive (i.e., from one level to another), as in Dell's (1986) system, not to a fully interactive system (i.e., one in which all levels are directly connected to each other). There are

three reasons for this choice. First, because our goal is to evaluate the minimalist position, we chose to contrast it with an alternative view that differs from it on the fewest assumptions and that is well represented in the existing production literature. Second, with respect to lexical processing, Rapp and Goldrick (2000) showed that error data from speakers with aphasia can be better accounted for by a locally interactive system. Finally, full connectivity is not a biologically plausible construal.

We consider minimal input and unidirectional flow of information separately because there exist hybrid models, which assume maximal input but a strictly feedforward flow of information, at least at certain joints in the system (e.g., Lloyd-Jones & Humphreys, 1997). Furthermore, as discussed in Boland and Cutler (1996; see also Norris, McQueen, & Cutler, 2000), in other domains of psycholinguistics, such as spoken word recognition, the crux of the modularity debate does not involve whether or not the system uses maximal input, which is generally agreed on, but, instead, whether or not information from one level can feed back and affect the processing at a previous level.

Our focus is on the processes concerning lexical retrieval and phrasal construction at the functional and positional level (see Figure 1). For clarity of exposition, we separately discuss lexical retrieval processes (the left side of the figure) and phrasal integration processes (the right side of the figure). The coordination of words and sentential frames is discussed as part of the phrasal integration processes.

Figure 1. Levels of integration in sentence production according to Garrett (1984). The figure emphasizes the levels between an intention the speaker wants to communicate to the phonological spell out of the sentence (positional-level processing). The processes beyond (i.e., phonetic-level processing, articulation) are not depicted. The arrows indicate the general flow of information, from the message to positional levels, without embedding any claim with respect to (a) whether at each level the input is minimal or maximal or (b) whether at each level feedback is assumed. V = verb; S = subject; N = noun.

Lexical Retrieval

There is broad consensus in the literature that word retrieval occurs in two stages. First, on the basis of the speaker's intended message, a meaning-based retrieval process is initiated (referred to as *lemma* retrieval in Figure 1, following Kempen & Huijbers, 1983, and Levelt, 1989), followed by a form-based retrieval process (*word-form* retrieval). Theories differ with respect to how lemmas are characterized. According to some authors (Butterworth, 1989; Garrett, 1982; Zorzi & Vigliocco, 1999), lemmas are lexical representations that specify the meaning of the word and that are organized into so-called "semantic fields" (Vinson & Vigliocco, 2002). However, according to Levelt et al. (1999), lemmas do not represent semantic information, but they do represent (or are strictly associated with) syntactic information, such as grammatical category, the type of phrase the word can be part of, the subcategorization frame for a verb, and other language-specific features (e.g., the grammatical gender of Italian nouns). Note that the assumption that lemmas (lexical entries) are syntactically characterized implies that there is not a complete separation between stored lexical information and syntactic encoding. Indeed many theories (both minimalist and maximalist) assume that production is lexically guided. Finally, other authors (Caramazza, 1997; Miozzo & Caramazza, 1997) have rejected the notion of an amodal lexical representation (lemma) altogether, assuming modality-specific (phonological and orthographic) lexical representations only.

Regardless of whether lemmas are considered to be semantically or only syntactically motivated, and regardless of whether these lexical representations can be considered as amodal, there is substantial agreement in distinguishing between a meaning-based retrieval process and a form-based retrieval process. This distinc-
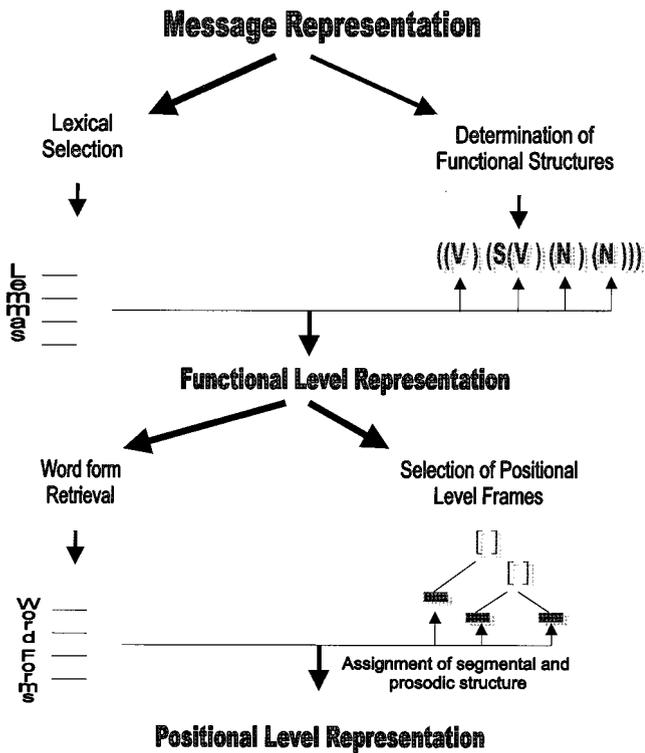
tion was originally motivated on the basis of the features of word substitution errors. Word substitution errors during spontaneous speech often show some relation between the target and the intruding word. According to Garrett (1975), two major types of word substitution errors can be found: substitutions in which the target and the intruding word have some meaning relation but no form similarity, as exemplified in (1, a and b), and substitutions in which the two elements have some form similarity but no meaning relation, as exemplified in (2, a and b).

(1) a. All I want is something for my *shoulders* [intended: elbows].
   b. I just like whipped cream and *mushrooms* [intended: strawberries].
   (from Garrett, 1976, p. 244)

(2) a. It is on my *soldier* [intended: shoulder].
   b. I've got whipped cream on my *mushroom* [intended: mustache].
   (from Garrett, 1993, p. 213)

Note also that both types of substitution errors obey a strict grammatical category constraint, that is, the error always has the same grammatical class as the intended word. We return to this issue later in the article.

Further evidence comes from experimental (reaction time) studies. A series of studies (e.g., Schriefers, Meyer, & Levelt, 1990) using primed picture naming (Glaser & Düngelhoff, 1984) showed that meaning- and form-related prime words can facilitate or interfere with the naming of a picture as compared with a neutral baseline condition. Of importance, these effects depend on the time interval between the presentation of the interfering word and the presentation of the picture. A meaning-related prime inhibited the retrieval of the target word if the prime word was presented before the picture, but a form-related prime facilitated the naming of the picture if the prime word was presented after the picture. These data are compatible with the lemma/word-form distinction and with an architecture in which lemma retrieval strictly precedes word-form retrieval, although they do not exclusively support such an architecture.

TOT states have also been interpreted as evidence compatible with the lemma/word-form distinction. These incidents are considered a failure in retrieving the word form after successful selection of the corresponding lemma by some authors (Butterworth, 1989, 1992; Garrett, 1984; Levelt, 1989) and as partial activation of lemmas and word forms by others (Dell, 1986; Meyer & Bock, 1992). Speakers in a TOT state can report, with a probability above chance, information about the word form—for example, the length in syllables and the word onset (A. S. Brown, 1991; R. Brown & McNeill, 1966). In addition, it has been shown that speakers in a TOT state can report syntactic information about the longed-for word. For example, Vigliocco, Antonini, and Garrett (1997) reported that Italian speakers in a TOT state can report the grammatical gender of the words they could not say with above chance probability (see also Miozzo & Caramazza, 1997). Furthermore, Vigliocco, Vinson, Martin, and Garrett (1999) reported that English speakers could report whether a noun was "count" or "mass." Finally, Iwasaki, Vigliocco, and Garrett (1998) extended the investigation to Japanese and to a grammatical class distinction (adjective vs. nominal adjective) with similar results.

Crucially, in all of these studies speakers were able to correctly report syntactic information even when they could not report any information about the word form. These findings are compatible with the claim that during the first step of lexical retrieval, semantic and syntactic information are accessible, whereas form-related information becomes accessible only during the second step.[1] These findings also indicate that lexical access involves retrieval of syntactic information along with meaning and form.

Evidence consistent with a two-step lexical retrieval process comes also from aphasia research. In particular, research has focused on speakers with *anomia*—an aphasic language disturbance in which speakers have particular difficulties in word retrieval. Similar to speakers without aphasia who are in a TOT state, speakers with anomia seem to be able to retrieve a lemma representation for the word, but they seem unable to retrieve the corresponding word-form information (Badecker, Miozzo, & Zanuttini, 1995; Goodglass, Kaplan, Weintraub, & Ackerman, 1976; Henaff Gonon, Bruckert, & Michel, 1989; Martin, Vigliocco, & Garrett, 1999). This is evidenced by the fact that speakers with anomia can generally report a good deal of semantic and syntactic information about the target word even when they can retrieve little or no phonological information.

Dissociations consistent with the two-stage theory have also been found in anomic naming performance. For example, Patient J.C.U. (Howard & Orchard-Lisle, 1984) and Patient E.S.T. (Kay & Ellis, 1987) seem to have distinct impairments, although both were diagnosed as having anomia. In particular, J.C.U. seems to have a semantically motivated lexical retrieval problem, but E.S.T. has a phonologically motivated lexical retrieval problem. This is evidenced by the fact that J.C.U. was helped by phonological cuing when experiencing word retrieval failures but was misled when the cuing was inappropriate for the target but appropriate for a close semantic coordinate—for example, it starts with "r" (racket) for the target *baseball*. E.S.T. was never misled in these circumstances. Further, J.C.U. showed poor performance in semantic tasks involving both comprehension and production, whereas E.S.T. did not show any semantic deficit in the same tasks. These two different patterns of performance can be explained in the general framework including lemmas and word-form representations as a problem in retrieving lemmas from conceptual descriptions (J.C.U.) and a problem in retrieving word-form information given successful lemma retrieval (E.S.T.). Similar to a speaker in

---

[1] Note that such a conclusion (namely, that syntactic information is available prior to phonological information) from TOT studies has been challenged. Caramazza and Miozzo (1997) reported that word-form retrieval is not dependent on syntactic retrieval in experiments on TOT induction. Hence, they concluded that word-form retrieval does not follow syntactic retrieval. The lack of statistical dependency between syntactic and word-form retrieval in TOT states has been replicated in other studies (Vigliocco et al., 1999). However, evidence for dependency comes also from other observations, as strong syntactic constraints apply to phonologically related word substitution errors (Vigliocco et al., 1999). Furthermore, in a series of simulation studies the lack of statistical dependency in a TOT situation was observed in models that assume strict dependency (Vinson & Vigliocco, 1999). On the basis of these observations, we maintain here a view in which syntactic information is available to the system prior to word-form information.

a TOT state, Patient E.S.T., but not Patient J.C.U., could provide information concerning word onset and number of syllables.

Furthermore, recently Foygel and Dell (2000) showed that an implemented dual-stage model of production can fit the performance of a variety of aphasic word naming errors, postulating disruption in a semantically motivated retrieval process, in a phonologically motivated retrieval process, or both.

Evidence for Minimalism

As the review above indicates, there is abundant evidence for the two-staged nature of lexical retrieval. Does the retrieval process entail minimal input and unidirectional flow of information? With respect to the semantically based retrieval process, there is general agreement that the process is maximalist and both maximal input and bidirectionality are allowed. For Levelt (1989), maximal flow of information at this level is allowed because lemmas are conceived of as shared between production and comprehension. Minimalism is assumed for the second step: namely, form retrieval.

In this section, the evidence compatible with discrete minimalist models is reviewed. In completely discrete models, minimal input and unidirectionality of processing ensure accuracy by blocking potentially interfering information. Efficiency is ensured by virtue of the speed of processing of such a modular architecture. For example, in Levelt et al. (1999), on the basis of the speaker's intentions, a lexical concept is activated and the corresponding lemma is selected. There is no transmission of activation to the word-form level before lemma selection.

Some of the studies reviewed above are also compatible with the discrete processing assumption. With respect to speech errors, the observation of only two main types of lexical substitution errors, those related in meaning and those related in form—see (1, a and

b) and (2, a and b)—is compatible with discrete models (e.g., Fromkin, 1971; Garrett, 1980) as is the dissociation between form and meaning observed in lexical retrieval failures (TOT states; e.g., Garrett, 1984).

With respect to reaction time studies, Schriefers et al. (1990) showed that semantic inhibition from a primed word was found only at stimulus onset asynchronies (SOAs) preceding the presentation of the target picture, whereas phonological facilitation was found only later. Most important, the authors found that the two effects did not overlap. Similar conclusions were reached by Levelt et al. (1991) with a modified task.

Phrasal Integration

Following a long tradition in production research, we begin our presentation of phrasal integration processes by using speech errors to motivate the postulated levels of integration. We then move to the discussion of a number of additional processing assumptions (restricted lexical guidance, incrementality, and internal self-monitoring) on which there is broad consensus in the literature. Finally, we review the evidence for minimalism for each of three joints in the system: (a) lexical retrieval and phrasal construction, (b) message-level planning and phrasal construction, and (c) functional-level and positional-level planning.

Levels of Integration

As for lexical retrieval, the core levels at which phrasal integration occurs have been inferred from properties of slips of the tongue. Particularly relevant here are *exchange errors*—errors in which two linguistic units switch position. Examples of exchange errors in English (Cutler, 1982; Stemberger, 1982) and Spanish (Del Viso, Igoa, & García-Albea, 1987) are reported in Table 1.

Table 1
Examples of Different Types of Exchange Errors in English and Spanish

Exchange type	Produced	Intended
Phrasal		
(Ia)	*Most cities* are true of *that*[a]	That is true of most cities
(Ib)	*las chicas* de *la cara* estan[b]	la cara de las chicas esta
	the-F, P girl-F, P of the-F, S face-F, S are	the-F, S face-F, S of the-F, P girl-F, P is
Stem—functional		
(IIa)	You *order*ed up *end*ing some fish dish[a]	. . . ended up ordering . . .
(IIb)	Pasame las *tortilla*s para la *patata*[b]	. . . las patatas para la tortilla
	Pass me the-F, P omelet-F, P for the-F, S potato-F, S	. . . the-F, P potato-F, P for the-F, S omelet-F, S
Ambiguous		
(IIIa)	She *sings* everything she *writes*[a]	She writes everything she sings
(IIIb)	le han dedicado *periodicos* sus *editoriales*[b]	le han dedicado editoriales sus periodicos
	to-her (they) have devoted periodicals-M, P on editorials-M, P	. . . editorials-M, P on periodicals-M, P
Stem—positional		
(IVa)	That's why they sell the *cheaps drink*[a]	. . . drinks cheap
(IVb)	me *despide deprimi*r a la gente en el tren[b]	me deprime despedir a la gente en el tren
	to-me (it) say goodbye-1P, S to depress to people on the train	to me (it) depress to say goodbye to the people on the train
Phonological		
(Va)	like a *l*ilting *w*illy[c]	. . . wilting lilly
(Vb)	me vas a *cas*ar	me vas a sacar una foto
	me (you) are going to marry	me (you) are going to take a picture

Note. Exchanged elements are indicated by use of italics. Spanish examples are followed by explanations. F = feminine; P = plural; S = singular; M = masculine; 1P = first person.
[a] Source of example is Stemberger (1982). [b] Source of example is Del Viso et al. (1987). [c] Source of example is Cutler (1982).

There are a number of important contrastive features for the different exchange errors, which motivated the hypothesis that sentence production proceeds via the development of a number of different planning frames.[2] The examples show that exchange errors involve different linguistic units. In *phrasal exchanges,* such as (Ia) and (Ib) in Table 1 two noun phrases (*most cities–that,* and *las chicas–la cara*) have exchanged. The fact that in (Ib) the inflectional morphology (number marking) has moved with the word stems indicates that the exchange has happened before the syntactic frame has been specified for number (e.g., that the phrase in subject position is singular and that the phrase in subject modifier or predicate position is plural). This invites the hypothesis that the error occurred during the process of assigning grammatical functions[3] to lemmas on the basis of the message. Consistent with that notion is the observation that the produced verbs agree with the number of the produced subject rather than with the intended subject (see Bock & Levelt, 1994, pp. 962–963, for a more detailed discussion).

In *stem exchanges* such as (IIa) and (IIb), two *stem morphemes* have exchanged (*order–end,* and *patata–tortilla*), whereas the inflectional morphology remained in place. Errors of this type are also referred to as *stranding exchanges* (Garrett, 1980). These errors contrast with phrasal exchanges, in which inflectional morphology moves with the stem. This suggests a different locus for the error. Stem exchanges arise after grammatical function assignment as errors in the insertion of a lemma within a functional-level frame.

The examples in (III) are *ambiguous,* because the words involved have the same number and gender. Therefore, the units potentially involved could be phrases, as in (I); word stems, as in (II); or fully inflected word forms. As discussed in Vigliocco and Zorzi (1999), the last possibility (an exchange of fully inflected forms), however, is very unlikely: In a richly inflected language such as Spanish, in which there are ample opportunities for observing exchanges of fully inflected forms, these cases were almost nonexistent (see Vigliocco & Zorzi, 1999, p. 58). Note that ambiguous cases are very common in poorly inflected languages such as English.

Errors such as (I)–(III) share characteristics such as the following: (a) There are grammatical category constraints—the words involved in the exchanges are of the same grammatical category (in 85% of the word exchanges in the MIT corpus; Garrett, 1980); (b) there is no meaning resemblance between the exchanged units, and we return to this point; (c) the units involved are part of different phrases (81%, in the MIT corpus), although they tend to be within the same clause (only 1% of the word exchange errors in the MIT corpus cross clause boundaries).

Examples of a different type of *stem exchange* are provided in (IV). These are also stranding errors; two stem morphemes are exchanged (*drink–cheap* and *deprim–despid–*). These exchanges contrast with the examples in (II) in their domain and characteristics. In exchanges such as (IV) the two exchanging units belong to the same phrase and do not share the same grammatical category, in contrast to the stem exchanges in (II). Furthermore, the exchanged units tend to share number of syllables and main stress (Garrett, 1980). In the framework presented in Figure 1, errors such as (IV) may arise because the wrong word forms are inserted in the positional-level representation.

Also arising during positional-level processing are *phoneme exchanges,* as in (V) /w/–/l/ and /k/–/s/. Errors such as (IV) and (V) have a number of characteristics in common such as the following: (a) There is phonological similarity between the units exchanged and (b) the units involved are either in two contiguous words or in two syllables within the same word, always within phrases. Furthermore, for sound exchanges, the source syllables are similar. These exchanges arise when the phonological makeup of the sentence is spelled out at the positional level.

Exchange errors contrast with substitution errors, as in Examples 1 and 2. Whereas substitution errors can be considered failures in retrieving stored lexical (or sublexical) information, exchange errors are cases of failure to integrate lexical and phrasal information. Buckingham (1979) and Garrett (1984) have argued that such a distinction is also supported by the observation that whereas individuals with Broca's and conduction aphasia are impaired predominantly in developing sentential frames either at the syntactic (Broca) or phonological (conduction) level, those with anomia and Wernicke's aphasia are predominantly impaired in lexical retrieval.

It is important to note here also that all the errors we have considered up to this point involved major category (*open class*) words and not minor category (*closed class*) words. The distinction between these two vocabulary types is well documented in the speech error literature (e.g., Garrett, 1975, 1980; Levelt, 1989), and it is a fundamental dimension in aphasic speech (e.g., Bradley, Garrett, & Zurif, 1980; Miceli & Caramazza, 1988). In sum, the contrasting features of errors such as (I)–(III) and (IV)–(V) are represented in theories of sentence production in terms of a distinction between functional-level processing, during which a hierarchically organized syntactic skeleton for the sentence is built, and positional-level processing, during which the serial ordering of sentential units is established and the intonational, metrical, and segmental makeup of the sentence is worked out (Garrett, 1984; see Figure 1). Errors such as (I)–(III) arise in this architecture in the following manner: Phrasal exchanges such as (I) arise when the selected lemmas are assigned the wrong grammatical function. Functional-level stem exchanges such as (II) arise because the wrong lemma is inserted in a correctly defined functional-level representation. Semantically related word substitutions such as in Example (1) also arise at this level as retrieval of the wrong lemma. The functional level representation is linearized during positional-level processing. During this stage, positional-level stem exchanges such as (IV) occur, when the word forms (or segments) are inserted in a wrong position in the frame. Form-

---

[2] Here, we depart from the traditional classification of exchange errors presented in Garrett (e.g., 1980) and use, instead, the classification proposed by Vigliocco and Zorzi (1999). The rationale for this departure is that when corpora are considered cross-linguistically, some error categories seem unsatisfactory. The main difference between the two classifications is in terms of what *word exchanges* (discussed in the text) are. Garrett (1980) proposed that word exchanges be considered as cases of wrong lemma insertion at the functional level, with stranding exchanges occurring at the positional level. Instead, we proposed that in fact word exchanges (meaning the exchange of two fully inflected word forms) do not exist. See the text for details.

[3] Note, however, that alternatively, phrasal exchanges may arise at the message level during the processes of thematic role assignment.

related word substitution errors such as in Example (2) also occur at this point. Sound exchange errors (V) occur when the segmental specification of a positional-level representation is spelled out.

We need to briefly digress at this point to clarify a potentially confusing use of the term *positional level* by different authors. This is necessary, because one of our purposes is to evaluate the separation between syntactic and phonological information. We consider here the positional level as morphophonological, following Garrett (1984) and Bock, Eberhard, Cutting, Meyer, and Schriefers (2001), and therefore this level overlaps with what is referred to by Levelt (1989) and Levelt et al. (1999) as *phonological encoding.* During positional-level processing, first morphemes are inserted into a prosodic frame that specifies the linear word order. Stem exchanges as in (IV, a and b) in Table 1 occur at this point. Then, phonemes are spelled out, and phoneme exchanges such as in (V, a and b) may occur. We do not maintain a separation between a level manipulating morphemes and a level manipulating phonemes because processing of both types of units seems to be affected by the same constraints: serial ordering and phonological similarity. Bock and Levelt (1994) instead used the term *positional-level processes* to refer to syntactic processes. In their description of *grammatical encoding,* the authors distinguish two stages devoted to building a syntactic representation for a to-be-uttered sentence: a functional level, in which grammatical functions are assigned to lemmas, and a positional level, in which a hierarchically and serially organized frame is built specifying closed class morphology (free and bound). In their view, positional-level processes are strictly syntactic, in contrast to our position. The main rationale for our choice is that there is no clear evidence supporting the notion of a positional-level representation on purely syntactic grounds, whereas there is evidence for a level of phrasal representation that, although strictly linked to syntactic phrasing, is prosodic in nature (F. Ferreira, 1993) and in which phonological similarity among units is important.

Further Processing Assumptions

Four further assumptions on language production are motivated on the basis of experimental evidence and error patterns. The first is referred to as *lexical guidance.* This is the notion that the encoding of a syntactic frame is, to an important extent, driven by lexical accessibility (Bock, 1987). In a series of studies Bock and colleagues (Bock, 1986a; Bock & Warren, 1985) showed that functional assignment depends on the *conceptual accessibility* of the particular nouns. Conceptual accessibility can differ because of inherent properties of the words (e.g., imageability), or as a result of semantic preactivation. In their studies, more accessible words tended to obtain more prominent grammatical functions, in agreement with the hierarchy of grammatical functions proposed by Keenan and Comrie (1977).

Lexical guidance, however, is not absolute. Our second assumption is that there is also a degree of independence between lexical retrieval processes and phrasal construction. As introduced above, the distinction between substitution errors and exchanges indicates a certain degree of separation between lexical and phrasal processing. Further support for this assumption comes from the finding of "*syntactic priming,*" first reported by Bock and colleagues for English (Bock, 1986b; Bock & Loebell, 1990; Bock, Loebell, & Morey, 1992) and reported in Dutch by Hartsuiker and colleagues

(Hartsuiker & Kolk, 1998a, 1998b; Hartsuiker, Kolk, & Huiskamp, 1999; Hartsuiker & Westenberg, 2000), that it is possible to prime abstract syntactic frames independent from the lexical content. For instance, Bock (1986b) showed this in a task where sentences were repeated and subsequently pictures were described. She found that syntactic structure tends to be persistent, that is, speakers often re-use a previously produced structure. This persistence effect could not be attributed to lexical priming, because the content words in prime–target pairs were always different. Furthermore, Bock and Loebell (1990) showed that the priming effect also persisted if closed class items differed, for instance, a *for*-dative (e.g., "The secretary baked a cake *for* the boss") primed a *to*-dative (e.g., "The undercover agent sold some cocaine *to* the rock star") as strongly as a *to*-dative prime.

It is important to note here that the assumption of a certain degree of independence between lexical and phrasal processes is not in conflict with the lexical guidance assumption (see also Garrett, 2000). Lexical guidance is an essential assumption of lexicalist theories of syntactic processing (e.g., Kempen, 1999; Pickering & Branigan, 1998; Vosse & Kempen, 2000, for production, and MacDonald, Pearlmutter, & Seidenberg, 1994, for comprehension). However, it does not constrain the degree to which syntactic information is specific to a given lexical element. A certain degree of independence between lexical and phrasal processes would run counter to lexical guidance only if one were to push the lexicalist view to an extreme (i.e., the position in which each lexical element would independently specify all syntactic properties).

A third processing assumption, broadly agreed on, concerns the *incremental* nature of sentence generation (e.g., V. S. Ferreira, 1996; Garrett, 1975; Griffin & Bock, 2000; Kempen & Hoenkamp, 1987). Incremental production implies that processes at a level $n + 1$ can start before the operations at a level n are concluded. For example, it implies that morphophonological processing can start before the syntactic encoding of a given clause is completed. Incremental processing would ensure fluent speech, hence efficiency. As argued before, different levels work on different unit sizes. If a representation at level n is partly completed but is sufficiently ready to send input for the encoding of a unit at level $n + 1$, this can still be viewed as totally compatible with a minimal input hypothesis. In our evaluation of minimal input we grant incremental processing and we are concerned with whether cascading of information (or additional nonnecessary information) is assumed beyond what would be expected on the basis of incrementality alone.

Finally, the fourth processing assumption is that speakers possess a self-monitoring system that allows them to inspect the well formedness of their speech and to take appropriate actions in the case that irregularities are detected. Self-monitoring is a particularly important concept for the present article. As we see below, data bearing on the question of whether there is feedback in the system can be reconciled with a minimalist framework, but only if a monitoring device is invoked. To set the stage for these arguments, let us briefly discuss monitoring systems. Self-monitoring refers to the processes with which a speaker inspects his or her speech plan for well formedness. These processes include detection of errors and other anomalies, implementing an interruption in speech, and planning a correction. An important distinction is that between monitoring systems that are localized within the speech

production system (Laver, 1980; Van Wijk & Kempen, 1987) or in systems that operate within a network that controls both production and comprehension (e.g., MacKay, 1992) and monitoring systems that make use of the speech comprehension system (e.g., Hartsuiker & Kolk, 2001; Levelt, 1983, 1989). So far, the available evidence has not allowed discrimination between perception-based theories, production-based theories, or theories that involve both perceptual and production processes. However, Hartsuiker and Kolk (2001), who provided a precise model of the time course of self-monitoring, concluded that a perception-based monitor is at least compatible with the timing evidence.

It is important to note that in some monitoring theories with a production-based component, such as MacKay's (1992) model, monitoring is subserved by feedback (see also Postma & Kolk, 1993, and Postma, 2000, for a discussion). Thus, with respect to those theories, the two alternative explanatory mechanisms are equivalent. This is not so with respect to Levelt's (1983, 1989) perceptual loop theory of monitoring. Let us briefly describe this theory, because here monitoring would be a distinct explanatory mechanism from feedback and it would, in many cases, rescue the notion of unidirectional flow of information. The perceptual loop theory is based on three main assumptions. The first assumption is that in addition to the speaker's self-generated speech (the outer loop), the monitor can also inspect speech before it is articulated (the inner loop). This assumption is uncontroversial: There is a plethora of evidence for prearticulatory monitoring (reviewed in Hartsuiker & Kolk, 2001). According to a study by Wheeldon and Levelt (1995), the representation of inner speech inspected by this loop is assumed to be a phonological code as it becomes available in real time.[4]

The second assumption of the perceptual loop theory of monitoring is that the system that inspects inner speech is the same system that also inspects overt speech: the speech perception system. This is a parsimonious assumption, because one would only need to postulate a single inspection system for both the inner and outer loop. However, there is little empirical support for this assumption, and a number of findings from the neuropsychological literature seem to argue against it (see Hartsuiker & Kolk, 2001). Notice that an implication of perceptual monitoring is that the inner loop is constrained by temporal limitations of speech perception. An open question is whether speech perception is sufficiently fast to account for (a) the existence of very short time intervals between the onset of an error in overt speech and the subsequent interruption and (b) the hypothesized existence of "covert repairs," that is, incidents in which speech is interrupted and corrected before the speaker has articulated the error. In support of Levelt's theory, Hartsuiker and Kolk (2001) showed that their computational model correctly predicted the distribution of intervals from error to interruption, in particular with respect to phonological and lexical errors (other kinds of errors, e.g., syntactic errors, hardly ever occurred in the error data that they studied).

The third assumption is that the monitoring system is localized within the conceptualizer (i.e., the processing component that generates the message). Both the inner and the outer loop would feed into the language comprehension system, and the output of that system (i.e., a semantic representation) is fed into the conceptualizer. Hence, the conceptualizer is the level at which the comparison between intention and realization occurs.

Evidence for Minimalism

Given the distinction between functional and positional level of integration in sentence encoding, as well as the distinction between lexical and phrasal processes, there are various degrees to which minimal cross talk among the different processes (i.e., information encapsulation) can be conceived of. The strongest view is the one proposed by Garrett (1975, 2000). Levelt et al. (1999) also proposed strong minimalism with respect to the joint between lexical retrieval processes and the processes engaged in phrasal construction. Garrett (2000) argued for independence at three major joints in the system; we discuss them in turn below. As in our discussion of lexical retrieval processes, for each of them the minimalist claim is that the shared information is minimal; that is, the putative encapsulated subsystems receive only minimal information in input and the flow of information is strictly unidirectional.

The Interface Between Lexical Retrieval Processes and the Processes Engaged in Phrasal Construction

With respect to this first joint, an example of a minimalist view is the model proposed by Levelt et al. (1999). Although this model is designed to deal with lexical retrieval processes and not with sentence-level processes, it is relevant because of its assumption that syntactic features (e.g., gender of nouns) are linked to lemmas. These syntactic features are retrieved only after lemma selection and would not affect the lexical retrieval process at any point. Thus there would be no phrasal construction for lemmas that are activated but not selected (minimal flow of information). Furthermore, a minimalist view predicts that frames under construction cannot affect lemma selection (unidirectional flow of information).[5]

To the best of our knowledge, no data support minimal input from lexical to phrasal processes. Some evidence compatible with the claim of unidirectional processing comes from recent work by Vigliocco, Lauer, Damian, and Levelt (2002). The authors reported an experiment in which bilingual Dutch–English speakers were asked to produce phrases in response to a prompt word in English. These prompt words were presented in the context (i.e., blocks) of other Dutch words. The context words were semantically related–unrelated, and/or had the same or different gender. A main effect of semantic similarity (i.e., a semantic interference effect) and a main effect of syntactic congruency (i.e., gender priming) were found in the Vigliocco, Lauer, et al. (2002) study. However, assuming bidirectional flow of information, we may also expect to observe an interaction between the two manipulations. This is because if activation feeds back from a syntactic frame that

---

[4] Note that this contrasts with an earlier proposal by Levelt (1989) according to which the code available to the inner monitoring channel was assumed to be a phonetic plan.

[5] Note that we limit our discussion here to syntactic properties that are lexically dependent such as gender of a noun. Such a feature in languages such as Dutch, German, Italian, French, and Spanish does not necessarily bear on the conceptual specification of the noun but is arguably an idiosyncratic lexical feature. Matters may be different if we consider other phrasal properties, such as argument structure of a verb (see our discussion above of lexical guidance and independence between lexical retrieval processes and phrasal construction).

is highly activated to the semantically driven retrieval process, it should enhance the activation of those lexical candidates that share the same syntactic features. For blocks of semantically related words, the feedback would enhance the activation of the target word as well as the activation of semantic competitors. Hence, the beneficial effect of sharing the same syntactic feature in this condition would translate into higher lexical competition and, therefore, slower response latencies. However, we found that the semantic and syntactic manipulations did not interact.

The Interface Between the Processes Responsible for Developing a Message-Level Representation and the Processes Engaged in Phrasal Construction (Functional-Level Processes)

What would constitute minimal input with respect to the message-/functional-level interface? According to Garrett (2000), "Semantic control is exercised in the initial stage of lexical and phrasal selection (i.e., lemma retrieval and grammatical function assignment above) but it is not directly implicated in the mechanisms of phrasal integration" (p. 45). As evidence for this claim, he argued that if there were semantic control on phrasal construction, units in exchange errors should exhibit semantic similarity (note, however, that Garrett, 1980, also argued that the lack of semantic similarity in word exchanges might be related to the low number of opportunities in the language to find such cases).

Additional evidence compatible with the claim of minimal input comes from studies of subject–verb agreement in sentence production. Agreement can be considered as a paradigmatic example of phrasal processing: the business only of those processes devoted to building abstract syntactic structures. These processes would be oblivious to conceptual and phonological information, and they would be affected only by syntactic information being computed during phrasal integration (in Fodor's, 1983, terms, *internal affairs*).

Consistent with the syntactic nature of the operations at this level are a number of experimental studies of agreement phenomena. These studies concerned agreement in number between the sentential subject and the verb, as well as agreement in gender (in Romance languages) between the sentential subject and a predicative adjective. In these studies, speakers were presented with sentence fragments as in Examples (3) and (4) and were asked to complete each fragment with a verb (or an adjective, in the studies on Romance languages reviewed below). Errors in the agreement between the subject and the verb, or between the subject and the adjective, were the dependent variables of interest.

(3)   The key to the cabinet

(4)   The key to the cabinets

With respect to syntactic manipulations, the following results have been reported. First, the presence of a mismatching noun in the subject noun phrase (NP), as in (4), has been shown to affect error incidence (e.g., Bock, 1995; Vigliocco & Nicol, 1998). We refer to this effect as the *effect of a mismatching modifier*. Of interest, the effect was larger when the mismatching noun was embedded in a prepositional phrase than when it was embedded in a clause, suggesting a role for syntactic structure (Bock & Cutting, 1992). Along similar lines, it has been shown that errors were not

just related to the linear proximity between the modifier and the verb: A similar pattern of results has been found both when speakers produced declarative sentences and when they produced questions (Vigliocco & Nicol, 1998) as in (5):

(5)   Is the key to the cabinets golden?

Furthermore, errors were also more common when the mismatching modifier was syntactically closer to the subject noun, as in (6) than when it was linearly close to the verb (Franck, Vigliocco, & Nicol, in press; Vigliocco, 1995), as in (7):

(6)   The helicopter for the flights over the canyon are ready.

(7)   The helicopter for the flight over the canyons are ready.

Finally, Hartsuiker, Antón-Méndez, and Van Zee (2001) showed that errors were more common when the mismatching constituent was a modifier of the subject phrase than when it was part of the direct object phrase, even though linear distance between verb and mismatching constituent was constant.

These studies indicate that agreement, being a syntactic relation, is sensitive to syntactic variables. The finding that agreement is affected by the hierarchical relations between the subject NP and the modifier, but not the linear proximity between the subject and the verb, and the finding of an effect of syntactic function lend further support to the hypothesis that agreement is computed during functional-level and not during positional-level processing.

Agreement, in minimalist accounts (Bock & Eberhard, 1993; Bock & Miller, 1991; Garrett, 2000), is not among the processes at the functional level under direct control from message-level structures. Therefore, it should not be affected by conceptual information. Bock and Miller, using the sentence-completion paradigm mentioned above, contrasted sentence fragments such as (8) and (9):

(8)   The baby on the blankets

(9)   The label on the bottles.

The example in (8) is a case of what they called a *single-token* preamble, and the example in (9) is a case of what they called a *multiple-token* preamble. For single-token preambles, the preferred interpretation implies only one instance of the referent of the head noun (in the example: one baby, sitting on a number of blankets). For multiple-token preambles, instead the preferred interpretation entails many instances of the head noun's referent (thus, many labels, each of which is attached to a different bottle). In other words, whereas (9) has a distributive reading, (8) does not. If the number of participants in the conceptual scene (one baby, many labels) is taken into account during phrasal construction, errors such as (11) should be more common than errors such as (10):

(10)   The baby on the blankets are crying.

(11)   The label on the bottles were peeled off.

This contrast is particularly interesting because the conceptual number (plural) of participants comes about only when the entire complex NP is considered, and in this respect it contrasts with the grammatical number of the subject NP (which is singular). If an effect of this contrast is found, it would be difficult to account for it in terms other than that the conceptual information has perme-

ated the processes of phrasal construction, introducing conflicting information.

Bock and Miller (1991), testing English, did not find a difference in error rates between these two conditions. This null effect was replicated, again in English, by Vigliocco, Butterworth, and Garrett (1996). Bock and Miller took their results as evidence in favor of insulation of syntactic operations: Information from the message does not affect internal affairs. Also consistent with the claim of minimal input from conceptual structures is Bock and Miller's finding that *animacy,* a factor that has been shown to affect the initial stages of sentence formulation (e.g., the decisions concerning grammatical function assignment; Bock & Warren, 1985; Comrie, 1978), did not affect the likelihood to observe agreement errors as would be expected on the assumption that conceptual factors do not permeate the internal operations of the grammatical encoder.

In sum, these findings support minimal input from conceptual structures to syntactic processes. What about the assumption of having only unidirectional flow of information between these two levels? This issue has received little attention and no experimental scrutiny in the literature. Therefore, we know of no data that support this assumption.

The Interface Between Processes at the Functional Level and the Morphophonological Processes Engaged at the Positional Level

In the architecture that serves as our framework (Garrett, 1984), the domain of syntactic processing is the functional level, and the domain of morphophonological processing is the positional level. This implies that the minimal input will be a hierarchical frame in which the lemmas have been inserted. However, other syntactic information, such as the syntactic category of a particular word, would constitute unnecessary input. One piece of evidence supports the minimal input hypothesis: the observation that stem exchanges such as (IV) in Table 1, although quite similar at the surface with errors such as in (II), are not constrained by syntactic variables, such as the grammatical category constraint.

Let us now turn to the question of unidirectional processing from the functional to the positional level. In support of unidirectional processing, Garrett (1993) reported that exchange errors such as (IV) and (V) are constrained by phonology, but no phonological constraints apply to exchange errors such as (I)–(III). This implies that positional-level information (i.e., phonological form) does not affect functional-level encoding. Furthermore, Bock (1986a) reported that priming the phonological form of a word does not affect the syntactic structure of a sentence.

This issue has also been investigated using agreement errors as the dependent variable. Indeed, agreement is well suited to test whether phonological information affects syntactic processing, because agreement is realized in languages through morphophonological regularities (e.g., in spoken English, the plural form of a noun and the third-person singular form of the verb is realized by adding /-s/ or /-z/ to the word stem). According to a minimalist view, we should not be able to find an effect of the morphophonological form on agreement errors. Bock and Eberhard (1993) addressed this question in an experiment in which they contrasted preambles such as (12) to (14):

(12) The beauty of the lace

(13) The beauty of the lake

(14) The beauty of the lakes.

The reasoning underlying the experiment is as follows: If the effect of a mismatching modifier is syntactic and the processes occurring during positional-level processing are separate from the processes occurring during the functional-level processing of the sentence, then errors should be triggered only when the modifier is syntactically marked as plural (14). A singular noun that happens to end in /s/ or /z/ (i.e., that has the same phonological ending as a plural), as in (12), would not trigger errors. This is what they found. They reported the same number of errors for the pseudoplural modifier condition (*lace*) and for the singular modifier condition (*lake*), and only the true plural modifier condition (*lakes*) induced high error rates. On the basis of this result, Bock and Eberhard (1993) concluded that only the syntactic specification of number affects agreement, a conclusion that is compatible with the minimalist view of unidirectional processing.

PROBLEMATIC DATA FOR A MINIMALIST LEVELS-OF-INTEGRATION VIEW

In the previous section we made a case for a levels-of-integration framework, and we presented the evidence for the minimalist assumptions of minimal input and unidirectionality with respect to this framework. The structure of the present section parallels that of the previous one. For both lexical retrieval and phrasal integration, we now discuss studies that have provided evidence against minimal input and against a strictly unidirectional flow of information from one level to another.

Lexical Retrieval

With respect to lexical retrieval, there is general consensus that the semantically driven retrieval process (from concepts to lemmas) uses maximal input and bidirectional flow of information. The controversial issue is whether form-based retrieval and phoneme retrieval are minimalist or maximalist. The two key assumptions of (a) minimal input and (b) strictly unidirectional flow of information of minimalist theories are discussed in turn below.

Minimal Input

With respect to lexical retrieval, minimal input implies that there is no cascading of activation in the retrieval of word forms. However, a number of studies have provided evidence that challenges the assumption of noncascading flow of information. First, Peterson and Savoy (1998), using a production priming procedure in which participants were presented with pictures to be named, and occasionally with words to be named, showed that, for near synonyms, both lexical candidates were activated to the level of word form. This was indicated by the finding of priming effects in producing *soda* after having named the picture *couch*. Thus, the near synonym *sofa* was also activated to the word-form level and hence facilitated the naming of the phonologically related word *soda.* However, this effect was restricted to near synonyms. No effect was observed for pairs that were only categorically related. Jescheniak and Schriefers (1998) also showed coactivation of near synonyms using the standard picture–word interference paradigm.

Can these findings be accommodated within a minimalist model? They can, but only at the cost of incorporating the additional assumption of multiple lemma selection in special circumstances, which can be considered as a special case of maximal input.[6] For instance, the experiments described above forced Levelt et al. (1999) to consider synonyms as a special case. They also admitted that there may be other special cases (e.g., hyponym–hyperonym relations such as *flower–rose;* Levelt et al., 1999, p. 17).

Second, Cutting and Ferreira (1999) reported data indicating that the cascading nature of activation is not limited to synonyms. Using the picture–word interference paradigm, these authors showed that picture naming for a homophonous word (e.g., *ball*) was speeded at early SOAs, when a distractor word semantically related to the nondepicted meaning (dance) was acoustically presented. This result supports the assumption of cascading activation beyond the special cases mentioned above.

Third, Griffin and Bock (1998) showed that naming a picture presented in sentential contexts (either semantically constraining or not) was faster in the constraining contexts. Crucially, the effect of context interacted with the frequency of the target name. In particular, a frequency effect was only observed when the sentence was nonconstraining. Because minimalist theories assume that word-form retrieval can only be accomplished after lemma selection, and because there is evidence that frequency effects arise at the word-form level (Jescheniak & Levelt, 1994), these results are, again, problematic for the minimalist framework.

Finally, starting from Starreveld and La Heij (1995), authors have further challenged the noncascading flow of information, using picture–word interference and distractors that were either semantically or phonologically related or both semantically and phonologically related to the target (Damian & Martin, 1999; Starreveld & La Heij, 1995, 1996).[7] The most important result from these studies is the finding of an interaction between semantic and phonological effects for a given (early) SOA. Minimalist views would predict, instead, additive effects of semantic and phonological relatedness.

These findings can also be accommodated in a minimalist theory (Levelt et al., 1999; Roelofs, Meyer, & Levelt, 1996) by introducing additional assumptions concerning the loci at which the visually or acoustically presented word can affect the retrieval of the name for a picture. In particular, the authors claimed that the attenuation of the semantic interference effect for semantically–phonologically related distractors arises as an indirect result of cascading in the language comprehension system. An acoustically presented distractor word would activate its word form in the input phonological lexicon, as well as a cohort of phonologically related words. Because there is cascading in comprehension, this will lead to the activation of the lemmas corresponding to this cohort of word forms. Lemmas, as opposed to word forms, would be shared between comprehension and production. Therefore, a semantically–phonologically related distractor word would activate both its own lemma and the lemma for the target word. A distractor that is only semantically related, however, would activate its own lemma, but not the lemma for the target word. Therefore, the difference in activation between the target lemma and the distractor lemma is larger when the distractor is both phonologically and semantically related than when it is only semantically related. Because the amount of semantic interference is a function of these activation

differences, a phonologically–semantically related distractor will yield a smaller interference effect. Note that this reasoning does not apply to distractors that are only phonologically related. A phonologically related distractor will boost the activation of the target lemma as well as that of its own lemma, but the distractor lemma is not part of the lemma's competitor set. Therefore, this activation increase for the target lemma is irrelevant with respect to the moment of selection. Phonological relatedness only exerts an effect at the word-form level. Hence, one should be able to observe a reduced semantic interference for mixed distractors at SOAs for which there is no phonological facilitation effect, and this is in agreement with Damian and Martin's (1999) data.

Note, however, that this explanation depends on many assumptions, some of which are well supported by data (i.e., the assumption of cascading in language comprehension; see Levelt et al., 1999, for a review), others of which are at least controversial (i.e., the assumption that semantic interference occurs only if the distractor is part of the competitor set; see Starreveld & La Heij, 1999, for discussion). Furthermore, it may be possible to explain these findings in a more parsimonious way on the assumption of a feedforward cascading account (see below).

Evidence problematic for the noncascading assumption held by minimalist theories comes also from investigations of slips of the tongue. Relevant here are blends, as in (15, a–c), and "mixed" substitution errors, as in (16, a–c). The examples are in English, Spanish, and Italian:

(15) a. didn't bother me in the *sleast* . . . slightest [least–slightest] (Boomer & Laver, 1968)
      b. a mí me gustan de ese *estipo* . . . de ese estilo [estilo–tipo] (Del Viso et al., 1987)
          (I like (those) of that *stype* . . . of that style [style–type])
      c. *indifendentemente* [indipendentemente–indifferentemente] ([independently–indifferently]) (Vigliocco, 2000).

In blends, two semantically close words (usually near synonyms) are both partially encoded in a single erroneous utterance, most often yielding a nonword. This implies concurrent activation of both lexical candidates up to the level of word-form encoding. Levelt et al. (1999, p. 17), in parallel to the synonym effects on reaction times discussed above, also discussed blends as an exceptional case of multiple lemma selection in which two (synonym) lemmas are simultaneously selected.

In mixed substitution errors (16), the intruding word bears both a semantic and a phonological relation with the target:

(16) a. *oyster* [lobster] (Garrett, 1993)
      b. para el próximo mes de *otoño* [Octubre] (Del Viso et al., 1987)
          (for the next month of autumn [October])
      c. vuoi una fetta di *limone?* [melone] (Vigliocco, 2000)
          (do you want a slice of lemon? [melon]).

---

[6] Technically, this would not be the same as allowing for cascading of activation because lemma selection would still precede word form activation. However, the distinction between the two becomes less and less clear.

[7] In these studies both visually (Damian & Martin, 1999; Starreveld & La Heij, 1995, 1996) and auditorily (Damian & Martin, 1999) presented distractors were used.

A prediction from minimalist theories is that mixed errors should not occur more often than chance would predict from the independent occurrence of semantically related and phonologically related substitution errors. However, Dell and Reich (1981), in an analysis of their Toronto corpus, reported a larger proportion of mixed errors than chance would predict. These authors estimated the occurrence of mixed errors on the basis of chance in two different ways that yielded highly similar results. The first estimate was based on randomly pairing all words (errors and targets) participating in the errors. The second estimate was based on phonological overlap between an independent set of semantically related pairs (i.e., near synonyms). Harley and MacAndrew (1995) reported a similar finding in their corpus. For Spanish a similar picture emerges, too (Del Viso, Igoa, & García-Albea, 1991).[8]

In addition to studies of spontaneously occurring speech errors, a number of authors have experimentally elicited speech errors in picture naming. For instance, Martin, Weisberg, and Saffran (1989) reported a higher than chance proportion of mixed errors in a study of picture naming (cf. Levelt & Maassen, 1981). This finding has also been replicated in picture-naming studies with speakers with language impairment (Dell, Schwartz, Martin, Saffran, & Gagnon, 1997; Martin, Gagnon, Schwartz, Dell, & Saffran, 1996).

In sum, as the review above indicates, there is abundant evidence for cascading of activation in lexical retrieval. Minimalist theories may accommodate these findings by incorporating additional assumptions, some of which (multiple selection in special cases) can be considered as a special case of maximal input.

Unidirectional-Only Processing

Let us now turn to the second key assumption: Do we need to postulate feedback from lower to higher levels? A number of studies have addressed this question with respect to the interface between lemmas and word forms, and, lower in the production flow, the interface between lexical representations and phonemes. We discuss these in turn. In principle, all of the studies discussed as compatible with maximal input are also compatible with the feedback hypothesis. However, these results can also be accounted for if we assume cascading of activation but maintain the assumption of unidirectionality. But, once again, the minimalist assumptions require additional assumptions. Let us consider two of these studies in some detail.

First, Damian and Martin (1999) interpreted their finding of an interaction between phonological and semantic codes as support for feedback between word forms and lemmas. However, this finding can be readily explained if we assume cascading of activation, but not feedback. The logic is as follows. In the mixed case and the semantically related case, the activation of the lemma for the distractor will be relatively high, compared with the phonologically related and the unrelated (baseline) case, because the semantically related distractors are in the competitor set for the target word. The high activation for the semantically related and mixed distractor lemmas results in more competition at the lemma level, yielding semantic interference. However, a further consequence of this high activation level is that in the mixed case, relatively more activation spreads to the word-form level than in the phonologically related case. As a result, the impact of the

phonological relation will be stronger, thus yielding the observed interaction.[9]

Second, mixed errors can also be explained without assuming feedback. The crucial question here is whether these errors are semantic errors that show a phonological overlap (which would be accounted for by assuming feedback) or whether instead they are phonological errors that show semantic overlap (which would be accounted for by just assuming cascading of activation). Rapp and Goldrick (2000), in fact, showed that mixed errors can arise in simulations in which only cascading of activation (and not feedback) was present.

It is interesting to note that Levelt et al. (1999) accounted for the mixed error effect in terms of erroneous selection of two lemmas (target and intruder). In their model when two lemmas are erroneously selected instead of one, the word-form encoding for the two will proceed in parallel. In the model it is more likely that the intruder may win the race when there is phonological overlap between the two than when there is not. Hence, multiple selection would account for mixed errors (Levelt et al., 1999, p. 35). Above we have discussed a minimalist multiple selection account for reaction time studies showing effects of synonyms and for blends (which also involve synonym words). Extending the notion of multiple selection from synonyms to semantically related words may account for the phenomenon of mixed errors, but it also raises the problem of identifying the special circumstances that trigger the erroneous multiple selection instead of the correct single selection, thus rendering this explanation ad hoc.

Finally, worth reporting here is the finding by Harley and Bown (1998) that the probability of experiencing a TOT state is higher for words in low-density phonological neighborhoods than for words in high-density neighborhoods. Although such an effect may be accounted for in terms of weaker word-form representation for words with few phonological neighbors (an account that, however, leaves unanswered why word forms with fewer neighbors should have weaker representations), a plausible alternative account of the findings is in terms of feedback. Lemmas corresponding to phonological forms that have no or few close neighbors receive little or no supporting activation from feedback from other related phonological forms.

Let us now turn to the interface between lexical representations and phonemes. Feedback between phonemes and lexical representations is supported by the finding of a *lexical bias* on sound (phoneme) errors (Baars, Motley, & MacKay, 1975; Dell, 1986; Dell & Reich, 1981), such that errors involving movement or substitution of phonemes result in existing words more often than chance would predict. The existence of this effect has been challenged by some authors. No lexical bias was reported by Garrett

---

[8] Note that Del Viso et al. (1991) concluded that there was not a significant effect of form on the semantically related errors in Spanish based on the exclusion from the analyzed set of all those mixed errors in which the degree of phonological overlap was "high" (see p. 175). In a previous analysis in which this subset of errors was included, a significant form effect was reported. We take this latter as a more representative analysis.

[9] This prediction does not depend on whether we assume that the distractor acts on both the lemma and the word-form level (as was argued by Levelt et al., 1999) or not.

(1976) for the MIT corpus or by Del Viso et al. (1991) in their Spanish corpus. However, an important issue is the determination of a chance baseline. Nooteboom (in press) recently analyzed a corpus of phonological substitution errors in monosyllabic words and proposed an improved method of establishing the baseline (based on the number of phonemes that are phonotactically legal at the compromised position and the proportion of these possible substitutions that would form real words). This analysis yielded clear evidence for a lexical bias effect: There were significantly more errors forming real words than would be expected by chance.

The lexical bias effect, however, need not be explained as a by-product of feedback. Baars et al. (1975) advocated an alternative explanation for the lexical bias effect. According to these authors, many errors can be intercepted, and covertly repaired, by the monitoring system (an *editing system,* in their terms). Of importance, the errors that form real words would be harder to detect for this putative system and thus would be intercepted less often than errors that form nonwords. A problem with this interpretation, however, is that Nooteboom (in press) did not find a lexical bias in overt self-corrections: Real-word errors were self-corrected as often as nonword errors. A monitoring account wherein the inner and outer channel are inspected by the same analysis system (such as the perceptual loop theory) would have predicted lexical biases both in the proportions of overt errors and in the proportion of overt self-corrections.

To summarize, in this section we presented evidence that is problematic for the assumption of minimal input and unidirectionality in lexical retrieval. Minimalist theories can account for the data that challenge minimum input, but again only if additional assumptions are made. Further, minimalist theories can account for the data that challenge unidirectionality, but only at the expense of assuming some form of maximal input and introducing further processing assumptions.

Phrasal Integration

The Interface Between Lexical Retrieval Processes and the Processes Engaged in Phrasal Construction

With respect to this interface and restricting our analysis to building phrases on the basis of lexically specified features (such as grammatical gender), a minimalist approach would entail that there is no frame-building for lemmas that are not selected. Thus, lemma selection ought to precede phrasal construction. The assumption of feedforward-only processing would imply that the processes that construct frames would not affect lemma selection.

Minimal Input

A conservative test of the minimal input assumption is an assessment of whether the retrieval of syntactic features that are not conceptually motivated (thus, features that depend on the lemma only) can occur before the selection process is completed. Two studies have addressed this question, both of which showed evidence for the hypothesis that syntactic features are retrieved prior to lemma selection. First, Marx (1999) showed that in spontaneously occurring semantically related noun substitution errors in German, the target and the intruder tended to share grammatical gender more often than chance would predict. Because semanti-

cally related substitution errors are errors in lemma selection, they can only be constrained by the grammatical gender of the target nouns (which are never selected) if the grammatical gender of the target noun is retrieved. Note here that semantically related substitution errors also tend to preserve the grammatical class of the target (the so-called "grammatical class constraint"). We believe, however, that the grammatical class constraint might arise for different reasons, most noteworthy the often greater semantic similarity between words belonging to the same grammatical class (see Vinson & Vigliocco, 2002, for a discussion). Hence we limit our discussion here to grammatical gender, because for this latter feature an account in terms of greater semantic similarity seems less plausible.

Second, Vigliocco, Vinson, Indefrey, and Levelt (2000) examined the same issue in an experimental situation. They induced semantically related substitutions for German nouns whose gender (masculine, feminine, or neuter) was not predictable on conceptual grounds. Speakers of German were asked to name pictures presented sequentially on a computer screen. In one condition, speakers were asked to provide a *bare noun* and in the other conditions, speakers were asked to provide a *noun phrase* (definite determiner + noun or indefinite determiner + noun). In regression analyses of the target intruder pairs, it was found that grammatical gender did not exert an effect above and beyond semantic similarity when bare nouns were produced; however, when phrases were produced, gender was found to be a significant predictor of the target–intruder pairing. These experimental results thus provide evidence against the assumption of minimal flow of information between lemma retrieval and phrasal construction. For the gender of the target to affect the intruding lemma, the syntactic properties of the target must be available to phrasal construction processes before lemma selection is completed.

Unidirectional Flow of Information

The results obtained by Marx (1999) and Vigliocco et al. (2000) also suggest that phrasal construction biases (via feedback) the lemma selection process. The fact that semantic substitution errors are constrained by grammatical gender shows, first, that the grammatical gender of nonselected items is active. But it also suggests that gender (a syntactic feature) influences the lexical retrieval process.

These results, however, do not necessarily falsify the assumption of unidirectionality. The logic is as follows. If we assume that phrasal construction starts before lemma selection is complete (maximal input), then in the case of a substitution error, the wrong lemma is initially inserted into the frame for the target lemma. This frame would thus contain an article that bears the gender specification for the target lemma rather than for the actually selected lemma. Let us further assume that the internal monitoring system can more easily spot and repair a substitution error in which article and noun have a different gender than an error in which they have the same gender. Then, substitution errors in which target and actual lemma have a different gender would more often be filtered out before articulation than errors in which both lemmas have the same gender, thus yielding a gender-preservation effect. Consistent with this putative mechanism is the finding that no gender-preservation effect was observed in a condition with indefinite nouns. The determiners for these nouns are invariant for the

genders tested (neuter and masculine). Thus, in that condition there was no overt gender-marking clue that the internal monitor could use in error detection.

The Interface Between the Processes Responsible for Developing a Message-Level Representation and the Processes Engaged in Phrasal Construction (Functional-Level Processes)

With respect to the assumption that the processes devoted to building abstract syntactic representations for a to-be-uttered sentence are insulated from conceptual information, minimal input would imply that conceptual information would be used only when it is necessary. For example, conceptual information is necessary to determine whether the subject head noun will be encoded as singular or plural. However, once this syntactic feature has been set, conceptual number information should not permeate further syntactic processes. A number of studies concerning agreement phenomena have provided results that challenge this strong assumption of minimal input. Below we review studies concerning the effect of conceptual (message-level) information on agreement processes. We first present data regarding cross-linguistic studies of subject–verb agreement and second present studies concerning subject–predicative adjective agreement in French and Italian. Finally, we discuss unidirectionality.

Minimal Input

Studies concerning number agreement. In the "A Minimalist Levels-of-Integration Approach to Sentence Production" section, we reviewed Bock and Miller's (1991) experiment that tested whether the number of tokens in the conceptual scene for a given sentence fragment affects the agreement process. In English, they did not find a difference in error rates between single-token and multiple-token conditions. This null result was replicated by Vigliocco, Butterworth, and Garrett (1996), also in English. Bock and Miller took the absence of such a "distributivity" effect as evidence for an insulation of syntactic operations.

However, our groups (Hartsuiker, Kolk, & Huinck, 1999; Vigliocco, Butterworth, & Garrett, 1996; Vigliocco, Butterworth, & Semenza, 1995; Vigliocco, Hartsuiker, Jarema, & Kolk, 1996) have extended the investigation of distributivity effects in agreement to other languages. In contrast to the results in English, we have found effects of distributivity in every other language we have tested: Dutch, French, Italian, and Spanish. The magnitude of the effect differed somewhat from language to language,[10] but the effect was reliably observed in each of them. These results disconfirm the view according to which universally the grammatical encoder is insulated from conceptual structures.

These results do raise the question of why no effect was observed in English. Vigliocco, Hartsuiker, et al. (1996) excluded an explanation in terms of differences between Romance and Germanic languages as the result was also found in Dutch, a Germanic language that is closely related to English. These authors were also able to dismiss a number of structural differences between languages that were likely candidates to account for the different results. In a recent article, Eberhard (1999) addressed the investigation of distributivity effects in English again, but she selected different materials. In particular, materials were selected so that

the items were rated equally high on "imageability" in both conditions. With this new set of materials, reliable effects of distributivity were also found in English. These findings suggest that the lack of an effect previously reported was due to a problematic selection of experimental materials. Whether a confound with imageability was really the culprit that resulted in a null effect in Bock and Miller (1991) and in Vigliocco, Butterworth, and Garrett (1996) remains to be determined.[11] What these results do clearly show is that it is possible to find effects of conceptual number cross-linguistically, a result that disconfirms the minimal input assumption.

In addition to distributivity, the potential impact of conceptual effects on number agreement has been investigated contrasting nouns referring to collectivities and nouns referring to individuals (Bock, Nicol, & Cutting, 1999; Vigliocco, 1995). In this regard it is interesting to note that for nouns referring to collectivities, two variants of the same language (British and U.S. English) have adopted different standards. British English allows the treatment of these nouns as either singular or plural (depending on whether the speaker intends to talk about the collectivity as a whole or whether the speaker wants to "separate" the individuals in the collectivity).[12] In U.S. English, instead, plural agreement with collective nouns is less acceptable. The same situation holds true (and even stronger) for Italian. In their investigation of U.S. English, Bock, Nicol, and Cutting (1999) reported that errors in subject–verb agreement for collective nouns were more common than for control nouns referring to individuals. Similar results were reported by Vigliocco (1995) for Italian.

Studies concerning gender agreement. Vigliocco and Franck (1999, 2001) have extended the investigation of conceptual effects on phrasal integration to a different form of agreement relation: gender agreement between the sentential subject and a predicative adjective in Italian and French. The properties of these two Romance languages allow us to establish whether conceptual effects on a syntactic operation such as agreement can be generalized among different agreement relations (and across languages). Of importance, these properties also allow us to explore a broader range of contrasts. In particular, we can contrast cases in which the conceptual information is absent to cases in which the conceptual information is congruent with the syntactic information, and we can also contrast cases in which the conceptual information is incongruent to cases in which it is congruent.

The syntactic number feature is retrieved from conceptual representations on the basis of a conceptual motivation. For example, number features for nouns usually (except for cases such as "mass"

---

[10] Differences in the methodological details, as well as differences in the stimuli used in the different languages, may be responsible for the differences in the magnitude of the effect.

[11] And we have good reasons to believe it is not. Eberhard claimed that the absence of a conceptual number effect came about because items in the single-token condition were less imageable than items in the multiple-token condition. Yet we found that imageability correlated positively with number of agreement errors regardless of condition.

[12] Bock, Humphreys, Butterfield, and Cutler (1999) presented data suggesting that in British English the preference to use a given collective noun as singular or plural does not necessarily depend on the speaker's intentions only but also depends on the frequency differences between the singular and plural form of that specific word.

nouns, such as knowledge, and *pluralia tantum* nouns such as *scissors*) depend on the number of entities the speaker wishes to discuss. What about gender? For this feature there is a clear distinction between conceptual and grammatical gender in Romance languages. In these languages, nouns are marked for either masculine (M) or feminine (F) gender. For nouns referring to animate entities, there is often a transparent relation between the gender of the noun and the sex of the referent (we refer to this situation as *conceptual gender*). Here, therefore, the gender feature of the noun can be assigned on a conceptual basis: whether the speaker wants to talk about a male or female entity.

However, for nouns referring to objects and abstract concepts, many animal names, and a number of nouns referring to humans, the genders of the nouns do not bear any relation to the sex of the referent (we refer to this situation as *grammatical gender*). For these nouns, gender is not a semantic property and thus cannot be assigned on the basis of the speaker's conceptual representation. Rather, it must be stored in the lexicon as an inherent property of the lemma (Vigliocco et al., 1997). Regardless of whether gender does or does not have conceptual correlates, agreement in gender between the subject and a predicative adjective is always required.

The question with respect to minimal input, then, is whether agreement features with a conceptual connotation (such as conceptual gender) have a different status than features with no conceptual connotation (such as grammatical gender). According to a minimalist view, in which agreement would only be controlled by syntactic properties of nouns, no difference should be found. This is because both nouns with grammatical gender and nouns with conceptual gender are marked as syntactically masculine or feminine. Using a constrained version of the sentence-completion task (Bock & Miller, 1991; Vigliocco et al., 1995), Vigliocco and Franck (1999) presented speakers of Italian and French with sentence fragments such as (17) and (18):

(17)   La ragazza nel parco [The-F girl in-the-M park-M]

(18)   La panchina nel parco [The-F bench-F in-the-M park-M].

Participants were required to complete the fragments with an adjective (that was provided). The subject noun in (17) has conceptual gender, but the subject noun in (18) has grammatical gender. The two nouns may receive the gender feature in distinct ways—on the basis of the sex of the referent in (17) and as part of the lexical representation in (18)—but if only the syntactic information is used for agreement, no difference in the proportions of errors should be found for the two cases. If, instead, the conceptual information permeates the agreement process, errors should be less common for items such as (17) than for items such as (18). This is so, because if conceptual information is allowed to influence phrasal integration and if it is redundant and congruent in (18), it should help correct agreement computation. Indeed, Vigliocco and Franck (1999) found conceptual gender effects. Furthermore, in a different study they excluded an account of these results in terms of animacy: Animacy per se did not account for the lower error rate in nouns with conceptual gender.

In another study, Vigliocco and Zilli (1999) asked whether the conceptual congruent specification of gender is used by speakers with Broca's aphasia. Two Italian-speaking patients were tested with the same materials that Vigliocco and Franck (1999) used. Overall, the 2 patients made significantly more errors than control speakers. But, crucially, errors were limited to the condition in which the subject noun had grammatical gender. In the condition in which the subject noun had conceptual gender, the difference between the patients and the controls was not significant. This shows that the patients were indeed using the conceptual information in their impaired syntactic-frame building (see Vigliocco & Zilli, 1999, for a discussion of how these results can be accounted for in a model of grammatical encoding).

Finally, Vigliocco and Franck (2001) performed a complementary test of conceptual influences on gender agreement. Italian and French provide us with the opportunity of testing the case in which the conceptual and the syntactic information about gender are conflictual, by looking at gender agreement between the subject of the sentence and a predicative adjective for subject nouns such as *vittima, victime* [victim-F] or *prodigio, prodige* [prodigy-M]. Nouns of this type have a fixed grammatical gender and can refer to either a male or female entity. Using such nouns, we can contrast cases in which the conceptual information is incongruent with the gender of the noun to cases in which the conceptual information is congruent. If the conceptual information does not influence agreement process, no difference should be found in the proportions of agreement errors for items with a match and items with a mismatch between the syntactic and the conceptual information. However, if conceptual information permeates the agreement process, one would expect more errors for items with a mismatch than for items with a match in grammatical and conceptual gender. In these experiments speakers were required to repeat and complete a sentence fragment with a (provided) predicative adjective. As a means of introducing the manipulation of the conceptual referent, speakers were first presented with a sentential context as follows:

Italian: Un camion ha investito Fabio–Fabiola che correva in bicicletta ascoltando musica
           [A truck hit Fabio–Fabiola who was riding the bike while listening to music]

French: Un camion a percuté Fabien–Fabienne qui roulait à vélo et ne l'avait pas entendu
          [A truck hit Fabien–Fabienne who was riding a bike and had not heard it].

Speakers were asked to read the context sentence silently, and subsequently they were presented with a sentence fragment such as (19) and requested to complete it:

(19)   a.   Italian: La vittima dello scontro
       b.   French: La victime de l'accident
            [The-F victim-F of-the-M crash-M].

In (19) the subject noun *vittima, victime* refers to a man (Fabio–Fabien) or to a woman (Fabiola–Fabienne). The sex of the referent is incongruent with the gender of the noun in the first case, but it is congruent in the latter case. In both languages, errors were consistently more common in the incongruent condition than in the congruent condition.

Furthermore, it is worth noting that in these experiments we manipulated the gender match and mismatch between the subject head noun and the modifying local noun. As discussed above, the effect of a mismatching modifier has been abundantly replicated across languages and agreement relations. However, in Vigliocco

and Franck's (2001) study, when the sentence fragment was introduced by a sentential context, no effect of a mismatching modifier was found in either language. Because such an effect surfaced in follow-up experiments in which no context was presented, the authors concluded that the (external) conceptual manipulation exerted an effect stronger that the (internal) syntactic manipulation.

It is also important to note here that these studies investigated agreement within clauses (subject–verb and subject–predicate), which is considered to be the domain of syntactic processes and, therefore, the strongest test of any claim of information encapsulation. Different forms of agreement are not considered to be the exclusive domain of syntactic processes. One example is subject–pronoun agreement. Languages differ in terms of whether this process is conceptually or syntactically driven (see Corbett, 1983, for a discussion). In the study discussed earlier concerning collective nouns in English, Bock, Nicol, and Cutting (1999) also reported data concerning tag pronouns and reflexives. In these cases, the proportion of plural verbs produced by speakers is extremely high. Along similar lines it is interesting to note that Italian and French nouns such as *victim* differ in what drives agreement in the subject–predicative adjective relation (grammatical gender of the noun) and in the subject–pronoun relation (either grammatical gender or sex of the referent; see Cacciari, Carreiras, & Barbolini-Cionini, 1997, for a discussion concerning parsing).

Unidirectional Flow of Information

What about the possibility of having bidirectional flow of information? With respect to the interface between the message level and phrasal construction, we know of no firm evidence that challenges the minimalist prediction of unidirectionality. However, we can raise a theoretical argument for some kind of bidirectionality. Note that minimalist theories allow for a form of indirect feedback, through the monitoring system. As an example of a monitoring explanation applied to the present interface, consider Levelt (1989). According to Levelt, the message level is re-inspected any time a syntactic revision is triggered during production in order to fulfill fluency requirements. Thus, the monitor would detect a lack of input and would trigger a restart, via the conceptual system. Levelt (1989) proposed that such a mechanism can account for word-form priming effects on phrasal structure (Bock, 1987).

However, we can also raise a theoretical argument for a direct form of feedback. This argument is based on the assumption that bidirectionality could have the function of ensuring that those aspects of syntactic form that need to be encoded in a language are in fact available in the conceptual representation developed by the speaker. As Roman Jakobson (1959) put it, "The true difference between languages is not in what may or may not be expressed but in what must or must not be conveyed by speakers" (p. 142). In fact, languages differ in what conceptual or formal properties need to be realized in sentential form. For example, the English word *friend* does not carry information about the sex of the friend. In Spanish the corresponding word is differentially inflected for a male friend *amigo* or a female friend *amiga.* Furthermore, in English, adjectives (e.g., *tall* in "The friend of Louis is tall") do not require gender agreement with the noun, but in Spanish they do (e.g., "El amigo de Luis es *alto*" or "La amiga de Luis es *alta*"). We therefore have an example of two languages that differ in

whether a certain type of conceptual information (sex) is obligatorily conveyed by the sentence (Spanish) or can be left unspecified (English). These facts are captured by the "thinking for speaking" hypothesis developed by Slobin (1996), according to which the information encoded in the message-level representation may differ among speakers of different languages. Empirical evidence that is compatible with this hypothesis has been presented by Slobin (1996).

Given this hypothesis, feedback from phrasal processes to conceptual processes may fine-tune the conceptual representation to what the phrasal processes require in the specific language one is speaking. Indeed, even Levelt (1989) granted this possibility, insofar as it applies to the language learner.

The Interface Between Processes at the Functional Level and the Morphophonological Processes Engaged at the Positional Level

Minimal Input

As mentioned before, in the framework that we have adopted, the minimal input to the positional level would be a hierarchical frame in which the lemmas have been inserted. However, the syntactic category of words is not part of this input. Consistent with this is the finding that positional-level stem exchanges such as (IV) in Table 1 do not adhere to the grammatical category constraint.

However, V. S. Ferreira and Humphreys (2001) provided evidence that is compatible with the view that syntactic category is accessible at the positional level. These authors induced stem exchanges by asking participants to repeat simple sentences as in (20) to (22) below:

(20)   They taped the record.

(21)   They framed the picture.

(22)   They shaved the beards.

The noun and the verb form of *record,* as in (20), are distinguished by a difference in stress (the ′rec-ord vs. to re′-cord). From time to time participants produced stem exchanges, saying, for example, "They recorded the tape." When an exchange was made, participants also changed the stress pattern of the words, changing therefore the grammatical category, so that it matched the phrase. Of importance, in this experiment exchanges were more common for cases like (20), in which participants changed the stress pattern, and (21), in which the noun and verb form are homophonic, than for unambiguous forms such as in (22). These results indicate that grammatical class information is used at the positional level, compatible with the hypothesis of maximal flow of information.

Can these results be accounted for in terms of minimal input? They can, but only if we assume that the frame at the positional level is marked for syntactic category and the morphemes (unspecified for grammatical class) would accommodate to the phrasal environment. This assumption, however, requires that positional-level phrasing is syntactic and not prosodic (F. Ferreira, 1993).

Hence, to accommodate the results, we should assume two positional levels: one syntactically specified and one phonologically specified. Such a (nonparsimonious) assumption would predict that positional-level stem exchanges—that is, exchanges

within phrases as in (IV) in Table 1—can come in two flavors: those that would occur during the building of a positional level in which the phrasing is syntactic and, hence, would not show phonological similarity effects and those that would occur during the building of a positional level in which the phrasing is prosodic and, hence, would show phonological similarity. Phonological similarity among exchanging morphemes in spontaneously occurring errors has been reported by various authors (Dell & Reich, 1981; Garrett, 1980), rendering an account in which only prosodic phrasing is assumed more plausible.

It should also be noted, however, that it is a controversial issue whether the grammatical category constraint is really present in spontaneously occurring morpheme stranding errors. There is large variation in the estimates of how frequently this constraint is present in corpora of spontaneous speech errors (41% in Garrett, 1980; 85% in Stemberger, 1985).[13]

Unidirectional Flow of Information

According to the strict unidirectional nature of the information flow, processes at the positional level should not affect processes at the functional level. However, Bock (1987) conducted a picture-naming study in which target pictures were accompanied by prime words that could be phonologically related to the names of elements in the picture. She observed that target words primed by the phonologically related word tended to appear later in the sentence, forcing changes in voice.

This finding can be accommodated in a minimalist framework, but in order to do so additional processing mechanisms need to be invoked. Levelt (1989) and Levelt and Maassen (1981) interpreted their results in the following manner. If a problem is encountered during the retrieval of the word form, a syntactic revision is initiated to fulfill fluency requirements. Thus, it would be an adjustment required under special circumstances, to maintain fluent speech.

Evidence compatible with a bidirectional flow of information between functional- and positional-level processing comes also from speech errors. Exchange errors occurring during functional-level processing—(I)–(III) in Table 1—should not show any phonological relation between the two exchanging words. However, Dell and Reich (1981) showed that phonological similarity between the exchanged words is greater than chance would predict.

Further evidence that challenges a minimalist view comes from studies of agreement phenomena. Here the question is whether manipulations at the positional and phonetic level affect processing at the functional level. Before we turn to these studies, it is essential that we introduce a few more general observations concerning the behavior of closed class words in speech errors. Some of these properties are essential to understand the levels at which the effects described below arise.

Closed class morphemes (both free and bound) are involved in a type of error that does not seem to occur for open class words: shifts.[14] Two examples in English are reported in (23) and (24):

(23)   what do you attribute *to* your longevity? [your longevity to] (Garrett, 1984)

(24)   that would be the same as add tenn*ing* [adding ten] (Garrett, 1984).

This finding has been interpreted by Garrett (1980, 1993) as indicating that closed class words are not retrieved during sentence production in the same way as open class words but that they are more strictly associated with the syntactic frames (for some additional evidence concerning the reading performance of patients with deep dyslexia, see Silverberg, Vigliocco, Insalaco, & Garrett, 1998).

Furthermore, closed class morphemes undergo processes of *accommodation* to errors involving the open class words to which they cliticize, such as in (25):

(25)   a meating arathon [an eating marathon] (Fromkin, 1971).

Crucially for our argument, this accommodation has been taken to indicate that closed class morphemes are represented in an abstract format until the phonetic level: the stage subsequent to positional-level processing. The reasoning is as follows: If the determiner in (25) was phonologically encoded during positional-level processing, it would not have accommodated to the sound exchange error, that is, to an error occurring at the same level. For the accommodation to take place, the sound exchange must precede the point in time in which the determiner's segmental shape is spelled out.

Getting back to agreement studies, a question that has been addressed in a number of languages is whether the morphological and phonological realization of agreement markers in the subject and local noun phrase affects accuracy of number and gender agreement on verbs and adjectives. As we did for conceptual effects on agreement, we present a summary of the major results, distinguishing studies that have addressed agreement in number between the subject of the sentence and a verb and studies that have investigated agreement in gender between the subject and a predicative adjective.

Studies concerning number agreement. Agreement is expressed in a number of languages through inflectional morphemes that specify the syntactic properties of nouns, determiners, verbs, and adjectives. Languages vary in their degree of morphological richness. For example, English is morphologically poor, because only number is consistently marked on nouns (with some exceptions, such as mass nouns) and even number is not always marked on verbs. Romance languages, on the other hand, are morphologically rich: Both number and gender are marked on nouns, verbs are marked for number in all persons and tenses, verbs are marked for gender in some occasions, and determiners and adjectives are always marked for number and gender. Some Germanic languages, such as German, explicitly mark case on nouns and determiners in addition to number and gender. The regularity with which the

_____

[13] In evaluating these differences, it is important to consider that English may not be the best test language to assess this claim. In fact, as we discussed above, when we consider a richly inflected language such as Spanish, it appears that true word exchanges do not exist and two types of stranding exchanges seem to come about: those that span phrasal boundaries and those that do not. The first type of stranding exchange is, in our opinion, what Garrett (1980) referred to as *word exchanges,* and the second type corresponds to *morpheme exchanges.*

[14] Garrett (e.g., 1984) also argued that closed class morphemes do not participate in substitutions and exchanges. With respect to phoneme exchanges, however, Dell (1990) showed that when frequency of use is controlled for, the probability of observing a phoneme exchange between open and closed class morphemes is not significantly different.

inflectional morphemes reflect syntactic features varies from language to language.

Variable among languages is also the amount and the transparency of phonological marking in NPs. For example, in Italian, determiners as well as nouns are marked for number and gender (with the exception of nouns starting with a vowel). In Dutch, there is also number marking and gender marking on the definite determiner. However, only for specific combinations of these features are gender and number transparently marked (when the noun is neuter and singular). This is different in German, where (for nominative and accusative case) gender and number are transparently marked on determiners for the singular (feminine, neuter, and masculine), whereas marking for number and gender is less transparent in dative and genitive case in the singular, and marking for gender is opaque for plurals of any case.

We reviewed above, among the evidence in favor of a strict separation of labor between functional and positional processes, the results reported by Bock and Eberhard (1993). The authors contrasted cases in which a modifier, ending with the phonemes /s/ or /z/, was denoting a real plural (lakes) or a singular noun (lace), and they reported that only real plurals induced agreement errors, a conclusion that is compatible with the minimalist view. Note, however, that in this experiment the subject noun was not manipulated; only the modifier was manipulated. A more stringent test of whether the phonological realization of the syntactic features has an impact on agreement must show that no morphophonological effect occurs when the subject noun is manipulated.

Vigliocco et al. (1995) conducted an experiment in Italian in which the phonological realization of the subject noun and the modifying noun was manipulated as illustrated in (26) to (29). In Italian, nouns are usually marked for number, singular (S) and plural (P), and gender in the word ending and, with respect to number, the pattern is highly regular: *-o, -a* is used for the singular, masculine and feminine, respectively; *-i, -e* is used for the plural, masculine and feminine, respectively. However, there is a class of nouns for which the same form is used for both the singular and the plural (0). Among these invariant nouns are very frequent words with main stress on the last syllable (e.g., *città,* city) and words borrowed from other languages that end in a consonant (e.g., *bar*):

(26) Il bar della borgata [the-S bar-0 of the-S suburb-S]

(27) Il bar della città [the-S bar-0 of the-S town-0]

(28) Il ristorante della borgata [the-S restaurant-S of the-F suburb-F]

(29) Il ristorante della città [the-S restaurant-S of the-F town-0].

These manipulations of the subject noun affected the rate of agreement errors. In particular, errors were more common for invariant nouns than for nouns explicitly marked. However, there was no effect of the manipulation of the modifying noun.

Hartsuiker et al. (2001) and Hartsuiker, Schriefers, Bock, and Kikstra (2000) manipulated the transparency of case and number marking in Dutch and German. In one experiment in Dutch, sentence fragments were presented that contained local pronouns as direct object. Dutch has two pronouns meaning *them: hen,* which is restricted to animate antecedents, and *ze,* which is used for inanimate antecedents. Crucially, *hen* is a word form that is unambiguously nonnominative and plural. On the other hand, *ze* is

a word form that is homophonous with the nominative forms *she* and *they.* Of interest, agreement errors were restricted to the condition with the case-ambiguous local pronoun *ze.*

Effects of case ambiguity were confirmed in studies with German. German prepositions in the local prepositional phrase differ in the case they require: either dative or accusative. Case is marked on the determiner that precedes the local noun. Crucially, for feminine and neuter nouns in the accusative, the determiner is homophonous with the determiner required for nominative case. Dative case is transparently marked, however. The data in two experiments with feminine and neuter local nouns revealed an effect of the mismatching constituent in the case-ambiguous, accusative condition. In the transparently marked dative condition, no such effect occurred.

Furthermore, both German and Dutch mark the number of the noun on the determiner, but the transparency of marking depends on the gender of the noun. In German, feminine nouns in the nominative require the determiner *die,* which is invariant for number. Neuter and masculine nouns, however, require the determiner *das* and *der* in the singular and *die* in the plural. Thus, for feminine nouns, number is ambiguously marked, and for neuter and masculine nouns, number is transparently marked. An experiment in German revealed that agreement errors in that language were mainly restricted to the number-ambiguous feminine subject condition. This was confirmed in Dutch, where errors occurred more often in the number-ambiguous common-gender subject condition than in the transparently number-marked neuter subject condition.

Studies concerning gender agreement. Gender agreement in Romance languages provides us with another excellent case for testing the role of morphophonological information on syntactic processes. In particular, Romance languages differ from each other with respect to how predictable of a given gender the realization of gender marking of nouns is. In a series of studies, Vigliocco, Franck, Antón-Méndez, and Collina (2001) investigated whether the morphophonological realization of gender has an impact on agreement accuracy in three languages: Italian, French, and Spanish.

For noun phrases in Italian, gender is predictably marked both on the noun's ending and on the determiner. Most nouns ending in *-o* are masculine, and most nouns ending in *-a* are feminine. There is then a class of nouns (about 15% according to our estimates) for which the same ending *-e* is used for both the masculine and the feminine (e.g., *ponte* [bridge-M], *febbre* [fever-F]). There are then irregularly ending nouns (about 5%), such as *mano* [hand-FS] or *fantasma* [ghost-M, S]. Gender is also predictable from the determiner. For all nouns starting with a consonant, the definite (and indefinite) determiner is different for the feminine (e.g., *la* casa [the-F house-F]) and the masculine (e.g., *il* tetto [the-M roof-M]). However, for the nouns starting with a vowel, because of vocalic neutralization, the same determiner is used for both genders (e.g., *l'antenna* [the-0 antenna-F], *l'elicottero* [the-0 helicopter-M]).

Spanish is relatively similar to Italian, but the degree of predictable mappings from word endings to gender is lower than in Italian (regular nouns ending in *-o* for the masculine and *-a* for the feminine constitute 68% of all nouns; Teschner, 1987). The determiner, as in Italian, is clearly marked for all nouns starting with a consonant (*el* for the masculine; *la* for the feminine). For nouns starting with a vowel, in contrast to Italian, Spanish also uses

gender-marked forms (*el* and *la*), except for a small set of nouns that start with the stressed word onset *a.* For these nouns, the masculine determiner is used (e.g., *el aqua* [the-M water-F]).

Finally, in French, gender is mapped in the morphophonological form of the noun in a more complex manner that allows for a high degree of unpredictability, although there are some regularities (Tucker, Lambert, & Rigault, 1977). In NPs, the determiner is marked for nouns starting with a consonant (*le* for the masculine, *la* for the feminine). For nouns starting with a vowel, however, the same determiner is used for both genders (e.g., *l'assiette* [the-S dish-FS]; *l'homme* [the-S man-MS]). Thus, for nouns starting with a vowel, French behaves like Italian.

In summary, Italian and French differ in the degree of regularity in the phonological marking of gender on the nouns. Spanish, instead, provides us with an additional interesting property: cases in which the determiner mismatches in gender with the noun because of phonological accommodation. In a series of experiments (Vigliocco et al., 2001) we have exploited these regularities to test whether the morphophonological realization of syntactic information has an impact on grammatical encoding. In a first series of two parallel experiments in Italian and French, we manipulated the degree of predictability of the subject noun's ending (predictable vs. unpredictable) and the markedness of the determiner (marked vs. unmarked) as in (30)–(33):

(30)   a. Italian: Il confine della zona
         [The-M border-0 of the-F region-F]
       b. French: Le cortège de la procession
         [The-M suite-0 for the parade]

(31)   a. Italian: Il premio della partita
         [The-M trophy-M for the-F game-F]
       b. French: Le chemin dans la vallée
         [The-M trail-M in the valley]

(32)   a. Italian: L'abete della foresta
         [The-0 pine tree-0 in the forest]
       b. French: L'exposé de la situation
         [The-0 description-0 of the situation]

(33)   a. Italian: L'aereo della compagnia
         [The-0 aircraft-M of the company]
       b. French: L'ancêtre de la trompette
         [The-0 ancestor-M of the trumpet].

The manipulation of the predictability of the noun's ending in French was based on published norms on the frequency with which each phonetic ending is associated with each gender (Tucker et al., 1977). In Italian the manipulation of markedness was simply realized by contrasting nouns that were transparently marked for gender (i.e., ending in *-o* or *-a*) with nouns that were unmarked (i.e., ending in *-e*).

We found that the nouns' endings played a role in agreement accuracy in both languages: Errors were more common for low-predictability nouns than for high-predictability nouns. However, the two languages differed with respect to the role of the marking on the determiner. Whereas this manipulation did not affect agreement accuracy in Italian, it had a large impact in French. This pattern of results is interesting for a number of reasons. First, we established that morphophonological manipulations affect pro-

cesses at the functional level. This parallels what we reported above for number agreement.

Second, the cross-linguistic difference that emerged is consistent with models of language use such as the competition model developed by Bates and MacWhinney (e.g., 1989), according to which the role of different cues (e.g., morphophonological marking) differs according to whether these cues are reliable and abundant or not. In Italian, the word ending is highly reliable and present in most cases, and this is the same for the determiner. However, in French, in most cases only the determiner provides a cue to gender. Hence, in French the cue validity of the determiner is higher than in Italian. The dependency of accuracy in computing agreement (which is a functional-level operation) on statistical regularities in the languages with respect to morphophonological markers suggests that this morphophonological information feeds back to functional-level processes.

Third, it is important to note that the effect of the determiner we observed in French strongly resembles the effect of accommodation of closed class words to a segmental error we described above—see (25). If we interpret the effect of the determiner along the lines of accommodation, then we have a case in which a phonetic-level process (i.e., selection of the form of the determiner on the basis of the phonological realization of the noun's onset) is affecting a functional-level process (i.e., agreement).

In a further series of two experiments in Spanish, we confirmed the role for the morphophonological ending of the subject noun. We also found that errors were far more common for those feminine nouns that require the masculine determiner (34) than for matched feminine nouns that use the feminine determiner (35), again demonstrating the effect of morphophonological cues on functional-level processing and creating problems for minimalist frameworks:

(34)   El agua en el pozo
       [The-M water-F in the well]

(35)   La vista del rio
       [The-F view-F of the river].

Do these studies provide unequivocal evidence for feedback from positional- (and even phonetic-) level processing to functional-level processing? Defendants of minimalist views can, again, invoke a monitor using the comprehension system to account for the data concerning both number and gender agreement. That is, the production system would produce equally often errors for noun phrases with transparent morphophonological cues and for noun phrases lacking morphophonological cues. However, the inner monitor would be more likely to detect and correct an error if morphophonological cues are present.

There are some problems with such a view, however. First, the monitoring theory, put forward by Levelt et al. (1999), on the basis of the results obtained by Wheeldon and Levelt (1995) cannot account for the results of the study concerning gender agreement in French. This is because the effect of morphophonology reported there can only be linked to the "inner" loop. In fact, the relevant errors were never overtly self-corrected. The inner loop in Levelt et al.'s (1999) theory would operate on phonological representations (i.e., the output of our positional-level processing). However, the spell out of the form of the determiner does not occur until phonetic-level processing (because it requires first the spell out of

the segmental onset of the noun as illustrated in the case of accommodation to error). Hence, if the inner monitor is assumed to operate at the positional level, it cannot detect and correct errors in the phonetic realization.

More generally, for a monitoring account of the results concerning both number and gender agreement, two strong assumptions regarding the power of the monitor must be made. First, the monitor should be fast enough to detect and repair syntactic errors before articulation. Second, the monitor performance has to be dependent on the morphophonological properties of the words in the different languages. Given Levelt's (1989) assumption that inner monitoring is carried out by the comprehension system, it follows that those variables that affect the monitor should also have an impact in comprehension. There is no direct evidence for either assumption.

With respect to the first assumption, Hartsuiker and Kolk (2001) reported simulations that showed that a perceptual-loop system is in principle fast enough to account for the time course of error detection and interruption of lexical and phonological errors. However, it is at least a possibility that this time course is different for syntactic errors. A reason for expecting such a difference is that the scope of syntactic errors, especially errors violating a long-distance dependency, is the clause rather than the word. A syntactic error becomes an error only because it is incompatible with syntactic aspects of words farther downstream or upstream in the clause. It is conceivable that detecting such an error is relatively time-consuming, as it requires the analysis of a larger unit of processing.

With respect to the second assumption, Bates, Devescovi, Pizzamiglia, D'Amico, and Hernandez (1995) and Taft and Meunier (1995) showed that the comprehension system is indeed sensitive to the degree of transparency in the morphophonological markers of gender in both Italian and French. In contrast, Vigliocco et al. (1995) reported a dissociation between production and comprehension, using the same materials that showed an effect of morphophonology in production—see (26)–(29). In the comprehension task, speakers were asked to indicate whether the subject of the sentence was singular or plural. In this latter task, an effect of the morphophonological marking of the subject noun was also observed, but in conditions different from the production task. That is, speakers were slower in deciding the number of the subject when the noun was unmarked but plural. In the production task, speakers made more errors, instead, for singular unmarked head nouns.

In sum, a monitoring account of the results reviewed in this section cannot be excluded. However, an account that assumes feedback has appeal because it does not require additional (and so far unsupported) assumptions.

Overall Summary and Implications

To summarize, we have presented a discussion of modularity in sentence production, with respect to an influential theoretical framework, Garrett's (1984) levels-of-integration architecture for sentence production, and with respect to lexical retrieval processes, as well as the interface between lexical retrieval and phrasal construction, we have extended this framework with detailed assumptions based on Levelt et al. (1999). We have supplemented this framework with four processing assumptions on which

there is a fair amount of consensus in the field: lexical guidance, (restricted) syntactic independence, incremental production, and the existence of an "inner" monitoring system.

The purpose of this article is not to question the multilevel nature of the architecture nor to question these processing assumptions. What is under debate is whether, given this theoretical framework, one can maintain the modularist viewpoint that there is information encapsulation for each of the different components of this architecture.

We have discussed four interfaces in the system assumed to bridge putatively separate modules. These interfaces bridge (a) the lemma/word-form level, (b) the lexical–functional level, (c) the conceptual–functional level, and (d) the functional–positional level. For each interface, we have discussed the two key assumptions of minimalism. The first is the assumption of minimal input. That is, nonnecessary information from an earlier level $n - 1$ cannot influence processing at a level n. The second assumption is the assumption of unidirectionality. That is, information from a later level $n + 1$ cannot influence processing at a level n via feedback. Below, we summarize the findings in terms of the complementary maximalist assumptions of maximal input and bidirectionality: That is, we discuss the evidence as support for these complementary assumptions rather than as falsifications of the minimalist assumptions.

With respect to lexical retrieval, we have considered the interface between lemmas and word forms. Our review indicates that the assumption of maximal input from lemmas to word forms is supported by a large body of evidence stemming from both reaction time studies (Cutting & Ferreira, 1999; Damian & Martin, 1999; Griffin & Bock, 1998; Jescheniak & Schriefers, 1998; Peterson & Savoy, 1998; Starreveld & La Heij, 1995, 1996) and error studies (Dell & Reich, 1981; Dell, Schwartz, et al., 1997; Martin et al., 1989; Martin et al., 1996; Del Viso et al., 1991; Harley, 1984, 1993; Stemberger, 1985). These results can also been accounted for within minimalist frameworks, but only at the expense of adding additional (ad hoc) assumptions (e.g., allowing multiple selection under special but broad circumstances). A number of these studies are also compatible with the assumption of feedback (e.g., Damian & Martin, 1999; Dell & Reich, 1981; Starreveld & La Heij, 1995, 1996) between word forms and lemmas. Support for the assumption of feedback comes also from further simulation studies by Rapp and Goldrick (2000). Rapp and Goldrick's goal was to assess whether both maximal input and bidirectional flow of information are necessary in order to account for different error distributions in patients with aphasia. Above we mentioned that a cascading model that does not incorporate feedback provides a sufficient account for mixed errors. However, a cascading architecture was found to be insufficient to provide a good fit to all patients discussed, whereas adding feedback to the same model produced a good fit. With additional assumptions, minimalist models can deal with the data, but only at the expense of granting a form of maximal input: multiple lemma selection both for synonym words and for semantically related words.

With respect to phrasal integration, we looked first at the interface between lexical retrieval and functional-level processing. Although very little work to date has addressed this interface, maximal input here is supported by evidence from syntactic constraints on semantically related substitution errors in German (Marx, 1999; Vigliocco et al., 2000). Proponents of the minimalist

framework might account for these findings, but again, only by incorporating additional assumptions (e.g., claiming that there is a qualitative difference in the information flow between erroneous and error-free performance). These data are also compatible with bidirectional flow of information, that is, phrasal effects on lexical selection. Accounting for the data without assuming bidirectional flow of information requires that additional assumptions be made. First, a form of cascading has to be granted (frame building for multiple lexical alternatives prior to lexical selection). Second, assumptions about the monitor have to made (i.e., that it prunes semantically inappropriate and syntactically ill-formed phrases more often than phrases that are only semantically inappropriate). On the other hand, the data of Vigliocco, Lauer, et al. (2002) seem to argue for unidirectional information flow at this particular interface, based on their additive effects of semantic interference and gender priming in picture naming. However, not only is it the case that "additivity does not uniquely support serial stage models; nonserial explanations of additive effects are sometimes possible" (Levelt et al., 1999, p. 8) but also such a conclusion is based on a single null effect that awaits replication with other methodologies.

Second, we looked into the interface between message-level and functional-level representations. Here the assumption of maximal input is supported by the finding of conceptual influences on the construction of number and gender agreement (Bock, Nicol, & Cutting, 1999; Eberhard, 1999; Hartsuiker, Kolk, & Huinck, 1999; Vigliocco, Butterworth, & Garrett, 1996; Vigliocco et al., 1995; Vigliocco & Franck, 1999, 2001; Vigliocco, Hartsuiker, et al., 1996; Vigliocco & Zilli, 1999). Within a minimalist framework, these findings can be accounted for, but again, only by assuming a qualitative difference between the processes engaged in erroneous and error-free performance. Support for the assumption of bidirectional flow here comes from the plausibility of the thinking for speaking hypothesis (Slobin, 1992, 1996), as well as the evidence supporting it. It is important to note here that even strong proponents of a minimalist view (Levelt et al., 1999) have assumed both cascading of activation and feedback with respect to the interface between concepts and lemmas.

Finally, we considered the interface between functional and positional level. Here, we also found evidence from speech errors that suggests maximal input (Dell & Reich, 1981; V. S. Ferreira & Humphreys, 2001), although, as we discussed above, these data are not as clear as the data we have reported for the other interfaces. Bidirectional flow of information is supported by the findings by Bock (1987) concerning effects of a phonological prime on functional assignment and the findings concerning morphophonological effects on agreement errors (Hartsuiker, Antón-Méndez, & Van Zee, 2001; Hartsuiker et al., 2000; Kuminiak & Badecker, 1998; Vigliocco et al., 1995, 2001; Vigliocco & Zilli, 1999).

Once more, the findings can be accounted for within minimalist frameworks, but only at the expense of incorporating additional assumptions about self-monitoring. In order to account for the results, two assumptions about the power of the monitor are required. First, it must be assumed that the monitoring system is sufficiently fast to detect and correct syntactic errors before articulation. Second, the probability for the monitor to detect the syntactic errors would have to be contingent on the morphophonological properties of the specific words in the phrase. To the best of our knowledge, there is no evidence for either of these two assumptions. The data on French reported by Vigliocco et al.

(2001) are particularly wounding for a monitoring explanation. In addition to the two assumptions mentioned above, a strong assumption would have to be made about the levels of speech planning to which the monitor has access (phonetic-level representations as well as phonological codes).

A MAXIMALIST LEVELS-OF-INTEGRATION APPROACH TO SENTENCE PRODUCTION

In the previous sections we have falsified a strict minimalist view of sentence production; hence, our goal here is to sketch in broad strokes a maximalist framework. The specific framework described below does not exhaust the range of possible nonminimalist positions, which range from highly structured and constrained (i.e., views in which levels of processing are maintained and in which necessary and nonnecessary information is differentially weighted) to strict maximalism (i.e., views in which levels of processing are not distinguished and in which all types of information are brought to bear on the processing).

Our proposal can be considered as a structured and constrained maximalist position. We assume multiple levels of integration, and we assume that the weight of different types of information differs depending on the specific level of integration and additional factors. Let us now discuss these two aspects in turn.

As outlined in the "A Minimalist Levels-of-Integration Approach to Sentence Production" section, our theoretical framework poses a major separation between functional and positional levels of representation for a sentence under construction. In this view, the functional-level processes are concerned with the mapping between speaker's intentions and a bound sentence-level frame that represents these intentions. That is, these processes realize a mapping between thinking and language. The domain of such a frame is semantic and syntactic. Positional-level processes are concerned with the mapping between a linguistic representation and a serially organized frame, that is, the mapping between a hierarchical frame and a linear frame. Such a mapping involves two steps: In a first step, word forms would be inserted in slots that correspond to linear positions; in a second step, segments would be linearized within phonological words. The domain of such a frame is prosodic.

Parallel to the distinction between functional- and positional-level frames, we posit two distinct lexical representations: lemmas, intended as semantically and syntactically specified representations (see Vinson & Vigliocco, 2002) which would guide the development of functional-level representations, and word forms, intended as representations that guide the unfolding of the syllabic and segmental spell out of words. For both phrasal construction and lexical retrieval, the main distinction is between a level guided by the message in which semantic and syntactic relationships determine the structure of the representation and a level in which a representation is specified for the linear order, timing, and phonological content. We further assume sublexical representations, such as phonemes, which would guide the unfolding of phonetic-level representations that would specify the actual timing of the articulatory gestures, although we do not discuss this aspect further. Hence, we posit the minimum number of levels that are motivated in terms of different functions served by the engaged processes and that can account for the data we presented in the previous sections. These architectural assumptions are not partic-

ularly new, in that they follow a number of other proposals in the field from Garrett (1975).

Each level is concerned with computing a specific type of representation (e.g., a hierarchically organized sentence frame at the functional level), and there are multiple sources of input available to those computations, but they have different weights. This situation is similar to the use of multiple sources of optical information to predict three-dimensional structure in visual perception. Such an analogy has been described by Kelly (1992). In this domain, for example, a primary source of information to the extraction of distance is motion parallax (in that there is a lawful relation between motion parallax and distance). However, it has been shown that viewers also use secondary probabilistic cues such as occlusion, relative size, and height in the projection plane (Bruno & Cutting, 1988).

The different weighting of information sources in sentence production is determined by the following factors. First, for each level there is some information that is primary and sufficient. For example, in retrieving a word form, the input from the corresponding lemma is primary and sufficient. As a more complex example, in computing number agreement (at the functional level), the primary and sufficient input for the agreement process (i.e., determining the number of the verb) is the syntactic number of the subject noun.

Second, the weightings differ with respect to the reliability of a given information source. Although the primary and sufficient source is always reliable, the degree to which a secondary source of information is relevant to a given level can vary across and within languages. For example, the degree to which morphophonological information is reliably associated with grammatical gender varies across Romance languages as we discussed in the "Problematic Data for a Minimalist Levels-of-Integration View" section. The degree to which semantic and phonological information is reliably associated with grammatical class is different in English (Kelly, 1992). It is interesting to note that reliability can apply at different levels of generality: from the language as a whole to a specific condition in one experiment (Dell, Reed, Adams, & Meyer, 2000). The general idea of different weighting of cues–constraints is shared with two other proposals in the psycholinguistic literature, namely, the competition model (e.g., Bates & MacWhinney, 1989; Bates et al., 1982) and the constraint satisfaction model (e.g., MacDonald et al., 1994; Trueswell, Tanenhaus, & Garnsey, 1994).[15]

Both modular and nonmodular architectures have to deal with the issue of how different information types are integrated. The modularity thesis provides a constrained framework for such integration. Along similar lines, some theories of depth perception propose that each cue to depth (e.g., motion parallax, stereo, texture, etc.) is computed independently and then integrated (bound) using a linear function (e.g., Landy, Maloney, Johnston, & Young, 1995). By assuming interactions among modules, additional constraints must be specified in order to avoid having an untestable model. Again, some theories of depth perception propose such interactivity among cues for depth, and it is only by specifying integration rules that the arbitrarily complex and untestable nature of the theories can be avoided (e.g., Nakayama & Shimojo, 1992).

In our version of maximalism, distinguishing among levels of integration, and assuming different weighting of the cues, we

attempt to render the proposal constrained. Can this work? The proposal by Kempen (1999) suggests it can. The model proposed by Kempen (see also Vosse & Kempen, 2000, for a parallel version dealing with comprehension) concerns syntactic encoding. It is silent with respect to how morphophonological features may come to play a role but presents a constrained proposal with respect to how conceptual features influence functional-level processes. Although it is far beyond the scope of the current article to describe the model, let us just consider some important aspects of it. Main assumptions underlying the model are as follows. First, the grammar is fully lexicalized—that is, lexical segments (which specify both categorical and functional information) are stored and constitute the building blocks of syntactic assembly. Second, assembly is governed by a single operation: unification of lexical frames (i.e., a probabilistic composition operation that combines lexical frames into larger units). Third, the likelihood of assembly (i.e., building a specific syntactic structure) depends on the level of activation of a given frame, the decay function as well as the strength of a given established unification link. The assembly process is, hence, dynamic, and competition between different possible assemblages governs the changes in the activation levels. In this view, conceptual factors can affect the system in two manners: acting on the strength of the unification links (i.e., syntactic structures that are supported by conceptual information have stronger links, and hence they inhibit more efficiently other structures) or directly on the activation of lexical frames. The author argued that such a model is successful in accounting for a variety of empirical phenomena concerning syntactic encoding. For our purposes what is most crucial is the fact that such a model presents a constrained example of how primary and sufficient information (lexical frames) is affected by secondary (conceptual) information.

Our proposal integrates both maximal input and bidirectional flow of information. Below we separately discuss these two aspects in turn. Our evaluation of minimalist and maximalist architectures has started by stressing the fact that theories of production have to account for the efficiency and accuracy of speaking, regardless of whether efficiency and accuracy are driving forces for a given architecture or, instead, they are a by-product of a given architecture. Hence, our assumptions of maximal input and bidirectional flow of information are discussed below with respect to whether, in fact, our proposed architecture can be considered as an efficient and accurate system. This seems even more important, given that a large amount of the evidence presented above as falsifying a minimalist architecture (and therefore compatible with

---

[15] Note that these other proposals have been developed to account not for sentence production but for language use more generally (the competition model) and for language comprehension (constraint-satisfaction models). We partly depart from these other models in the following manner. In contrast with both alternatives, we maintain a levels-of-integration approach, which we justify with respect to the "task demands" of production (in contrast to comprehension, e.g.), as discussed in the section "The Relation Between Comprehension and Production." Furthermore, we depart from the notion of cue availability (and hence cue validity) as presented in the competition model. The reason for this latter departure is empirical; in Vigliocco et al. (2001) we found effects of morphophonological marking of gender in French for a type of cue with extremely low availability but high reliability.

a maximalist architecture) concerned cases in which fluent and accurate production breaks down.

Maximal Input

As illustrated in the "Problematic Data for a Minimalist Levels-of-Integration View" section, there is a vast amount of evidence compatible with the claim of maximal input, but what could be the potential benefits of having maximal input from one level to the next? As Norris et al. (2000) put it with respect to spoken word recognition, "The task of recognizing speech is surely so difficult that one might expect the human speech recognition system to have evolved to take full advantage of all possible evidence that would be of assistance in recognizing spoken words" (p. 324). Surely, producing sentences is not an easier task. Therefore, using all evidence may prove to be beneficial in this domain as well.

Maximal input in production aids efficiency at three of the four joints we considered, with the potential exception of the interface between lemma and word-form retrieval. In fact, with respect to this interface, cascading of activation from lemmas to phonological codes can preactivate the phonological codes and, hence, speed up the retrieval process. However, cascading of activation would also activate other word forms, which might result in competition at this level. It is an empirical question to establish whether the benefits of preactivation may exceed the costs of activating competitors.

With respect to the interface between lexical retrieval and phrasal integration, maximal input implies that phrasal construction can occur before lemma selection, that is, not only for the eventually selected lemma but also for competitor lemmas that are highly activated. This can be beneficial in terms of efficiency because it implies that building phrases does not have to wait for the retrieval process to be completed. However, as for the interface between lemma and word-form retrieval, there may also be costs associated with activating additional frames. In contrast to the interface between lemma and word-form retrieval, with respect to this interface there is empirical evidence suggesting that either there are no costs associated with building phrasal frames for nontarget lemmas or at least the costs associated with them do not exceed the benefits. V. S. Ferreira (1996) reported that speakers can generate syntactic phrases in parallel, without competition between them. Syntactic phrases generated in parallel would instead "race" each other.

With respect to the interface between message-level and phrasal integration, maximal input implies that conceptual information that is not necessary can still be used. Vigliocco and Franck (1999) discussed one way in which this can aid efficiency in phrasal integration processing. Let us consider, for example, the processing of number agreement between the subject and the verb. The conceptual number is used to assign the appropriate syntactic number to the subject noun. This is the necessary information. However, the conceptual number could also be used beyond the assignment to the noun to mark the verb independently. Note that verb marking does not require conceptual information, because it can be achieved on the basis of the marking of the noun. This nonnecessary information, however, may be advantageous with respect to efficiency because it allows some degree of morphophonological encoding for the verb even before the integration with the noun phrase is achieved. Finally, with respect to the interface

between functional- and positional-level processing, maximal input in terms of cascading of activation, again, can aid efficiency in that it allows for the greater possible degree of incrementality.

Does maximal input aid accuracy? Maximal input does not seem to aid accuracy during lexical retrieval or at the interface between lexical retrieval and phrasal integration. For both of these two joints, allowing cascading of activation from multiple lexical candidates, if anything, may hinder accuracy!

The situation is different for the other joints. At the interface between message-level and functional-level representations, maximal input can aid accuracy. As discussed by Vigliocco and Franck (1999, 2001), because the conceptual information is convergent with the syntactic information in most cases (i.e., cases in which the conceptual and the syntactic information about number and gender are discrepant are arguably extremely rare across languages; hence, in most cases the conceptual and syntactic information are convergent), it can be used when the syntactic information is lost. An argument along similar lines can be made for the interface between functional-level and position-level representations. Here, again, in most cases there is a high degree of correspondence between information at the functional level (e.g., a plural syntactic marker) and at the positional level (e.g., the inflectional morpheme -s) in English.

In summary, maximal input by itself does not provide us with a system fully efficient and accurate. In particular, maximal input by itself may even favor certain types of derailment of the system, namely, lexical retrieval errors and errors in the integration of lexical entries in phrasal frames. However, maximal input is well supported by the reviewed data and, hence, has to be incorporated in a model of sentence production.

Bidirectional flow of information, on the other hand, is not as well supported by the available data. In particular we have noted in the previous section that a number of error findings reported in the literature as compatible with bidirectional flow of information may also be accounted for in terms of maximal input and/or by assuming an error monitor. Does this mean that we need to abandon the notion of feedback altogether and propose a cascading feedforward model of sentence production in a manner analogous to what has been done in auditory word recognition by Norris et al. (2000)? We believe not. Below we first present two general arguments for the utility of the notion of feedback; we then present how feedback can, in fact, aid both efficiency and accuracy by acting as the basis on which an inner monitor can work. Combining bidirectional flow and maximal input, the resulting system can fulfill the two fundamental requirements of efficiency and accuracy.

Bidirectional Flow of Information

Among the general reasons for assuming feedback, prominent are the fact that feedback can play an important role in language development and the fact that feedback is biologically plausible.

With respect to language development, there are some logical arguments favoring feedback, at least at some of the joints we have considered. For instance, with respect to the interface between the message and the functional level, we have introduced the notion of thinking for speaking: The messages prepared by speakers of different languages have to be tuned to the properties of the languages themselves. For the learning child this tuning can be

achieved by an interaction between the message and the functional level. This would allow the specification at the message level of those conceptual features that have to be grammaticalized in a given language. The importance of such a tuning has been recognized also by authors who subscribe to a minimalist view (Levelt, 1989).

If feedback is used by the language-learning child, then what about the "mature" adult language processing system? Levelt (1989) wrote, "Although conceptualizing and grammatical encoding *are* interacting for the language acquiring child, the mature speaker has learned what to encode when preparing a message for expression . . . . In short the system has become autonomous" (p. 105; see also Norris et al., 2000, for a similar argument regarding auditory word recognition). Hence, although the language-learning child may use feedback to establish the mature state of the system, feedback subsequently disappears. This view implies a discontinuity between learning a language and using a language. Such a claim seems to us to be less plausible and less parsimonious than assuming continuity between the learning and using of a language, because it implies that a mechanism in place at a point in time then disappears (cf. Seidenberg & McDonald, 1999). One function for adult processing of bidirectional flow of information is to allow continuous implicit learning. More generally, bidirectional flow of information can be used in learning for those interfaces where the mappings between one type of representation and another are predictable to some extent. Take the mapping between lemmas and word forms. This mapping is traditionally considered as prototypically arbitrary. However, some regularities exist. For example, across languages morphological composition sets a correspondence between the domain of meaning and the domain of form. Even for morphologically simple words, there is some evidence that the correlation between phonological form and meaning is not completely arbitrary (Shillcock, 2000). Also at the word level, the mapping between syntactic features and form is not arbitrary. Kelly (1992) reported that English speakers use information about main stress location on disyllabic words in order to infer grammatical class (i.e., they are sensitive to the fact that nouns tend to be stressed on the first syllable and verbs tend to be stressed on the second syllable). In Italian, the correlation between grammatical gender of a word and its word ending is also a reliable cue, and speakers are sensitive to it (see Bates et al., 1995). The fact that speakers are sensitive to these correlations suggests a role for them in learning the language as argued by a number of researchers (Bates & McWhinney, 1982; Kelly, 1992; Saffran, Aslin, & Newport, 1996). Speakers' sensitivity to such correlations does not stop after childhood as illustrated in research using, for example, artificially imposed regular and irregular mappings between syntactic and phonological markers in artificial grammar experiments (e.g., Brooks, Braine, Catalano, Brody, & Sudhalter, 1993; Frigo & McDonald, 1998) or artificially imposed phonological constraints (Dell et al., 2000).

Another reason for postulating feedback is the fact that a large body of evidence suggests that the brain uses feedback for various aspects of perception and action. Feedback is a general principle of connectivity in the brain. As reported by Spitzer (1999), only 0.1% of pyramidal cells in the cortex are concerned with input and output operations. The remaining 99.9% are internal connections. Given the sheer number of neurons and given the fact that a pyramidal neuron has on average about 10,000 output connections,

a simple calculation shows that on average a signal is back to where it started after three synaptic transmissions.

Feedback connections seem to have various roles in neural processing. In the visual domain they can filter the input and improve its quality by changing the sensitivity of the afferent pathways to certain aspects of stimulation (e.g., Alonso, Cuideiro, Perez, Gonzalez, & Acuna, 1993). They may also be crucial in the synchronization of adjacent populations of neurons in the cortex, considered to be the neural mechanism of feature binding (e.g., Finkel & Edelman, 1989). Obviously, there is not necessarily a direct correspondence between cognitive systems and their neural implementation. However, we believe that psychological models should be at least consistent with biological constraints (Grossberg, 2001). Biologically plausible computational models of perceptual processing (e.g., Grossberg, 1976) use feedback to achieve fundamental goals such as the stable development of connections and adult learning. According to Cutler and Norris (1999), there would be a distinct disadvantage to having feedback in perceptual processing (speech recognition): A listener would be so much influenced by top-down expectations that he or she would not be able to accurately perceive, for example, a speech error. On the other hand, according to Grossberg (2001), feedback plays an integral role in overcoming perceptual uncertainties while at the same time being insufficient to force us to see things that are not there.

What are the consequences of assuming bidirectional flow of information for accuracy and efficiency? There are various arguments that suggest an important role for feedback in learning, but does feedback help the adult language user in achieving the goals of efficiency and accuracy? According to Postma (2000), feedback subserves three possible functions in human performance: a tuning function (i.e., so as to adjust future action plans as during learning), a directive function (controlling the action on-line), and a corrective function (i.e., monitoring). We have already discussed the tuning function of feedback above. However, if feedback were also to serve directive and corrective functions in language production, it would directly contribute to efficiency and accuracy.

With respect to a putative directive function of feedback in production, Postma (2000) cited Dell's (1986) model of lexical access. In that model, feedback from sublexical units to lexical representations increases the activation level of the lexical target. This would diminish the selection probability of incorrect sublexical units. Combining the advantage of cascading of activation (preactivation of target units at a subsequent level) and the directive function of feedback, the resulting mechanism is both efficient and accurate.

As a further example of a directive function of feedback, Dell, Schwartz, et al. (1997) suggested that feedback from phonological codes to lemmas could serve the function of a readiness signal that would indicate at the lemma level whether the corresponding phonological code is ready for encoding. Such a line of argument can be generalized to the other joints in the system we considered, and in this manner, again, feedback would aid in ensuring efficient and accurate performance.

Can feedback have a corrective function? We have pointed out in numerous places that the error data reviewed in the "Problematic Data for a Minimalist Levels-of-Integration View" section do not unequivocally support a feedback account if a monitoring system, via the comprehension system (Levelt, 1983, 1989), is

assumed. Surely, monitoring plays an important role in language production, and it is very plausible that we can monitor both our own overt speech and others' speech using the language comprehension system. Furthermore, there is a fair amount of evidence for the existence of both an inner and an outer monitor (see Hartsuiker & Kolk, 2001, for a review of the literature and simulations that support this notion). However, one should be careful in evoking a monitoring device any time the data seem to argue for feedback. Each time this is done, additional assumptions have to be made about the power of the monitoring system and about its error-detection performance. Without any independent confirmation of all these assumptions, monitoring explanations remain ad hoc. In the remainder of this section, we present three further arguments against such explanations. First, there are no data that convincingly show that inner monitoring proceeds through the comprehension system. Second, inner monitoring through the perceptual loop runs into a serious logical problem: the fact that the perception system has to keep track of both the internal and the external speech signal. Third, proposals in the literature that assume production-internal monitoring systems are often based on the existence of feedback. However, if that is true, monitoring can no longer be viewed as an alternative explanation to feedback—instead, it should be considered as one of the raisons d'être for feedback!

Levelt et al. (1999), when invoking monitoring explanations, presented them as a form of indirect feedback. It is indirect because it is mediated through the language comprehension system. The "indirectness" here rests on the assumption of a perceptual inner loop. However, there is no empirical evidence that unequivocally supports the notion that inner monitoring proceeds through language comprehension (Hartsuiker & Kolk, 2001).

Furthermore, attractive though this notion may be, there is an important problem associated with it. Although inner and outer monitoring through comprehension have traditionally been assumed to proceed by different "channels," the two channels in fact converge. This poses a serious coordination problem: The comprehension system needs to deal with two phonetic codes that represent the same message and that are only slightly asynchronous. How can the comprehension system distinguish between the two codes, and how does it prevent interference between them?[16] If these two codes are processed with a slight delay, why do we not experience the consequences of this delay (e.g., "echoes"). Why do we not sometimes mix up the codes, so that we would experience later words earlier?

In the literature on monitoring, a number of alternative proposals to perception-based monitoring have been made (see Postma, 2000, for an overview). These proposals either assume production-internal monitoring mechanisms (e.g., Berg, 1986; Laver, 1980; Postma, 2000; Postma & Kolk, 1993; Van Wijk & Kempen, 1987) or monitors that are based on a network that partially subserves both production and comprehension (MacKay, 1992). A commonality between a number of these proposals is that the mechanism for monitoring is based on feedback (e.g., Postma & Kolk, 1993). In particular, a code that is selected at a given level will receive feedback from codes at a lower level. The amount of feedback received is indicative of whether the correct codes have been selected or not. This information can then be used by a (local) monitoring mechanism in order to detect derailments in processing and subsequently start corrective measures.

An (inner) monitoring mechanism outside the language comprehension system that uses feedback is attractive for at least two reasons. First, it solves the coordination problem alluded to above. Because inner and outer monitoring would use separate channels, they do not interfere with each other. Second, we assume continuity between language learning and mature language use. Using feedback for monitoring processes would be a way to make use of an information stream that would be present anyway but no longer subserves the function of learning to the same extent as in the child. An interesting implication of feedback-based production monitoring is that feedback helps accuracy, but indirectly. A direct consequence of feedback is an enhancement of activation at higher processing levels, which may be used in directive control. However, another consequence of feedback is that it provides information about the activation dynamics at lower levels (see, e.g., Berg, 1986; Botvinick, Braver, Barch, Carter, & Cohen, 2001; Mattson & Baars, 1992; Schade & Laubenstein, 1993). Thus, feedback can inform a higher level whether processing at the level below proceeds according to plan, it can signal discrepancies, and it can trigger corrections.

The discussion above does not imply that monitoring is only executed by a production-internal system. Speakers can listen to their own overt speech, and it is quite conceivable that they use the comprehension system for monitoring external speech. What about internal speech? We do not dispute that monitoring via the comprehension system may occur when people "listen" to their prearticulatory speech (i.e., when they speak subvocally). Hence, the comprehension system might be involved in monitoring of both overt and internal speech. This is, however, different from the characterization of the inner loop provided by Levelt et al. (1999) and Wheeldon and Levelt (1995). According to these authors, the representation being monitored internally is the phonological code as it unfolds in time and, crucially, not a phonetic code.[17] Their main argument is that Wheeldon and Levelt's (1995) participants were able to perform an inner monitoring task while speaking out loud (which is known to suppress phonetic representations). However, this type of monitoring could not have been inner speech in the sense of "talking in our heads," or to put it in Levelt's (1989) terms, as "involving a certain degree of consciousness" (p. 12), because "level-internal" representations, such as the phonological code, are generally viewed as being inaccessible to consciousness (Bock, 1982; Levelt, 1989). Only end products, such as a phonetic plan, can become conscious, provided that the speaker attends to them (Levelt, 1989; see also Postma, 2000).

Our view on these matters is straightforward. Inner speech exists: It is phonetic in nature, we are consciously aware of it, and we can inspect it through self-perception. However, it does not continue when we speak aloud, because of articulatory suppression. There is, at the same time, inner monitoring when we speak aloud, but this type of monitoring is internal to the production system; has access to, inter alia, phonological codes; and does not

---

[16] Hartsuiker and Kolk offered a speculation—a so-called *parse buffer*—that would deal with situations of overload to the parsing system (e.g., situations in which the parser receives simultaneous input from the inner and the outer loop).

[17] Note that, therefore, they deviated from earlier proposals by Levelt (1983, 1989).

lead to conscious awareness. In support of this view, there is both anecdotal evidence for self-correcting without awareness and empirical evidence compatible with it (reviewed in Postma, 2000). Hence, monitoring via the comprehension system of both overt and internal speech in our proposal does not run into the coordination problem simply because listening to our inner speech is suppressed when overt speech is produced. To summarize, we are proposing here that feedback may be conceived as at the core of an inner monitor, along the lines previously proposed by Postma and Kolk (1993) and other authors. Such an inner monitoring system does not exclude the possibility of monitoring our speech via comprehension, but we impose a restriction: Perception can monitor either our overt speech or our silent speech, but not both of them at the same time.

Having established the plausibility of assuming feedback in general, and having presented general arguments for the utility of assuming feedback at the service of efficiency and accuracy, we still need to reevaluate whether, incorporating feedback, the system can be considered as efficient and accurate specifically for those interfaces in which we saw that maximal input by itself fails to fulfill these two functions. We believe that for each of those interfaces, adding feedback, given its directive and corrective functions, would ensure efficiency and accuracy at all joints. We have already discussed how feedback would solve the problem of assuming only maximal input at the interface between lemmas and word forms. As discussed in Dell, Schwartz, et al. (1997), feedback would bias lemma access so that lemmas with retrievable forms are selected, hence preventing speakers from being lost in TOT blockages. This advantage may well outweigh the cost involved in activating additional word forms by virtue of cascading of activation: Cascading of activation would prime the corresponding word form, and feedback would provide information about its retrievability, hence aiding efficiency in the form of maintaining fluency and aiding accuracy in the form of preventing retrieval blockages. At the joint between lexical retrieval and phrasal integration, we can apply the same reasoning.

THE RELATION BETWEEN COMPREHENSION AND PRODUCTION

Language production and language comprehension are the two counterparts of adult language use, and although they are different because their input (a thought for production and a sequence of sounds for comprehension) and their goal are different (expressing and establishing "who did what to whom"), they must be similar, and they must be interconnected. The idea that the comprehension system may aid accuracy in production acting as a monitoring system (Levelt, 1989, and our revised hypothesis presented here) assumes similarities and interconnections between the two systems.

Issues of minimal input and of unidirectionality of the flow of information have been a driving force for research in the comprehension literature as much as, or even to a larger extent than, in the production literature. Boland and Cutler (1996) presented an enlightening discussion concerning minimal input in the comprehension literature. These authors proposed that the crucial question with respect to modularity of processing is not whether multiple candidates are generated at a given level and passed to a subsequent level (which they referred to as *multiple output*) but whether or not higher level processes influence directly the generation of candidates at a prior level. In other words, they argued that models in which maximal input is allowed can be considered as modular (in the parsing literature such an example is the incremental interactive theory developed by Crain & Steedman, 1985, and Altmann & Steedman, 1988), whereas models allowing for bidirectional flow of information are not (constraint-satisfaction lexicalist views such as those proposed by MacDonald et al., 1994, and Trueswell et al., 1994, belong to this class). Of interest, they pointed out that parsing models allowing for multiple outputs– maximal input (such as in Altmann & Steedman, 1988) have been traditionally considered as interactive models in contrast to autonomous models in which a single syntactic structure is proposed for further semantic processing (Frazier, 1987). This situation is similar in the language production literature, in which so-called "truly" autonomous models are models that entail only minimal input.

In our review of language production studies, we have stressed the fact that maximal input is largely supported by empirical evidence while at the same time unambiguous evidence for bidirectional flow of information is scanty. Is the situation analogous with respect to sentence comprehension? We believe not. Work conducted within a constraint-satisfaction framework indicating a role for probabilistic biases in syntactic analysis strongly suggests a system that uses information from different levels as it becomes available. Models in which syntax would dispose of what is generated by statistical perceptual strategies that would provide an initial semantic analysis (i.e., the semantic analysis would precede syntactic analysis that would come into play only when semantic analysis on its own provides ambiguous results) have also been proposed (Bever, Sanz, & Townsend, 1998).

If production and comprehension are similar and interrelated, why then would there be more evidence compatible with bidirectional flow of information in comprehension than production (besides the fact that more people work in comprehension than production)? One possibility could be that the degree of interactivity may differ in comprehension and production. As mentioned above, production and comprehension have different goals. Speaking involves fulfilling the communicative intention by realizing the details of the form: syntactic, morphological, phonological, and articulatory details. Details at each and every of these levels have to be spelled out by the system for each sentence. In Merrill Garrett's words, "The production system must get the details of the form right in every instance, whether those details are germane to sentence meaning or not" (Garrett, 1980, p. 216). The comprehension system, instead, has as its main objective to work out an interpretation of a given input. The details of the form are necessary in this enterprise; however, the degree of full specification that is required for these details, because of the different goal of the two systems, may not be the same.

In production, the system has to take into account dependencies from one type of information to another to a larger extent than in comprehension. For example, getting back to agreement, in order to correctly realize the gender marking on an adjective (*ross-o* [red-M] or *ross-a* [red-F]) in a sentence such as (36), the system needs (a) the gender of the noun, which being lexically specified implies that the noun *panchina* [bench-F] has been retrieved, and (b) to have established that *rossa* is a predicate of *panchina*. If one or the other is missing, the phonological form of the inflection *-a* cannot be realized. Some of these interdependencies are language

specific: Gender marking is one example, and word order constraints is another.

(36)   La panchina nel parco e' rossa.

Johnson-Laird (1983) proposed the following metaphor for comprehension (as a variant of Occam's razor):

> If you can read the signposts, you don't need a map. Parsing a sentence is, indeed, like following a set of signposts: you can reach your destination without having to keep a record of the directions in which they point. (p. 67)

This metaphor may fit comprehension, the listeners being the persons reading the signposts; however, the speakers are in charge of putting the signposts in place, for the listeners to use them. Placing the signposts cannot be achieved unless the speakers know where the signposts should go, which implies that the speakers need to use a map!

Hence, because of the differences in their goals, production and comprehension may require different degrees of temporal overlap of different processes and different processing strategies. These different requirements, however, do not imply that the systems' architecture is different. In fact, we argue that assuming the same architecture is parsimonious. Bidirectional flow of information, we argue, is essential to achieve stable learning. In adult sentence comprehension, then, it may serve functions such as reducing ambiguity at each level; in production, instead, bidirectional flow of information may serve the function of ensuring accuracy, aiding a production-internal monitoring system.

CONCLUSIONS AND FUTURE DIRECTIONS

We presented a review of the sentence production literature assessing how the flow of information is regulated from one level of processing to a subsequent one. We selectively focused our discussion on four major joints in the production system: the interface between meaning-based and form-based retrieval, the interface between lexical retrieval and phrasal construction, the interface between conceptual and syntactic encoding, and finally the interface between syntactic and morphophonological encoding. For each interface, we argue that minimalist views that incorporate the assumption both of minimal input and of unidirectional flow of information cannot account for the available data, unless additional strong assumptions are made. Hence, maintaining a levels-of-integration architecture, we propose that language production uses both maximal input and feedback in the service of efficiency and accuracy. Maximal input would allow fluent and efficient performance, and feedback would serve accuracy by subserving a production-based monitoring system.

Such a proposal opens the way to new investigations. Having established that nonnecessary information plays a role is only a first step; the next questions concern identifying in which conditions nonnecessary information plays a role and the relative weight of different types of nonnecessary information in different conditions. The work conducted within the competition model provides a first example of a research agenda along these lines; a vast literature in visual perception, concerned with investigating the relative strength of different cues in different conditions, provides another example.

References

Alonso, J. M., Cuideiro, J., Perez, R., Gonzalez, F., & Acuna, C. (1993). Influence of Layer 5 of Area 18 of the cat visual cortex on responses of cells in Layer 5 of Area 17 to stimuli of high velocity. *Experimental Brain Research, 93,* 363–366.

Altmann, G. T. M., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition, 30,* 191–238.

Baars, B. J., Motley, M. T., & MacKay, D. G. (1975). Output editing for lexical status in artificially elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior, 14,* 382–391.

Badecker, W., Miozzo, M., & Zanuttinni, R. (1995). The two-stage model of lexical retrieval: Evidence from a case of anomia with selective preservation of grammatical gender. *Cognition, 57,* 193–216.

Bates, E., Devescovi, A., Pizzamiglio, L., D'Amico, S., & Hernandez, A. (1995). Gender and lexical access in Italian. *Perception and Psychophysics, 57,* 847–862.

Bates, E., & MacWhinney, B. (1982). Functionalist approaches to grammar. In E. Wanner & L. R. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 173–218). Cambridge, England: Cambridge University Press.

Bates, E., & MacWhinney, B. (1989). Functionalism and the competition model. In B. MacWhinney & E. Bates (Eds.), *The crosslinguistic study of sentence processing* (pp. 10–75). Cambridge, England: Cambridge University Press.

Bates, E., McNew, S., MacWhinney, B., Devescovi, A., & Smith, S. (1982). Functional constraints on sentence processing: A cross-linguistic study. *Cognition, 11,* 245–299.

Berg, T. (1986). The aftermath of error occurrence: Psycholinguistic evidence from cutoffs. *Language and Communication, 6,* 195–213.

Bever, T. G., Sanz, M., & Townsend, D. J. (1998). The emperor's psycholinguistics. *Journal of Psycholinguistic Research, 27,* 261–284.

Bock, J. K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review, 89,* 1–47.

Bock, J. K. (1986a). Meaning, sound, and syntax: Lexical priming in sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 575–586.

Bock, J. K. (1986b). Syntactic persistence in language production. *Cognitive Psychology, 18,* 355–387.

Bock, J. K. (1987). An effect of the accessibility of word forms on sentence structures. *Journal of Memory and Language, 26,* 119–137.

Bock, J. K. (1991). A sketchbook of production problems. *Journal of Psycholinguistic Research, 20,* 141–160.

Bock, J. K. (1995). Producing agreement. *Current Directions in Psychological Science, 4,* 56–61.

Bock, J. K., & Cutting, J. C. (1992). Regulating mental energy: Performance units in language production. *Journal of Memory and Language, 31,* 99–127.

Bock, J. K., & Eberhard, K. M. (1993). Meaning, sound and syntax in English number agreement. *Language and Cognitive Processes, 8,* 57–99.

Bock, J. K., Eberhard, K. M., Cutting, J. C., Meyer, A. S., & Schriefers, H. (2001). Some attractions of verb agreement. *Cognitive Psychology, 43,* 83–128.

Bock, J. K., Humphreys, R., Butterfield, S., & Cutler, A. (1999, September). *Collective agreement in British and American English.* Paper presented at the Architectures and Mechanisms of Language Processing Conference, Edinburgh, Scotland.

Bock, J. K., & Levelt, W. J. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945–984). San Diego, CA: Academic Press.

Bock, J. K., & Loebell, H. (1990). Framing sentences. *Cognition, 35,* 1–39.

Bock, J. K., Loebell, H., & Morey, R. (1992). From conceptual roles to

structural relations: Bridging the syntactic cleft. *Psychological Review, 99,* 150–171.

Bock, J. K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology, 23,* 45–93.

Bock, J. K., Nicol, J., & Cutting, J. C. (1999). The ties that bind: Creating number agreement in speech. *Journal of Memory and Language, 40,* 330–346.

Bock, J. K., & Warren, R. K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition, 21,* 47–67.

Boland, J. E., & Cutler, A. (1996). Interaction with autonomy: Multiple output models and the inadequacy of the great divide. *Cognition, 58,* 309–320.

Boomer, D. S., & Laver, J. D. M. (1968). Slips of the tongue. *British Journal of Disorders of Communication, 3,* 2–12.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108,* 624–652.

Bradley, D., Garrett, M. F., & Zurif, E. (1980). Syntactic deficits in Broca's aphasia. In D. Caplan (Ed.), *Biological studies of mental processes* (pp. 269–286). Cambridge, MA: MIT Press.

Brooks, P. B., Braine, M. D. S., Catalano, L., Brody, R. E., & Sudhalter, V. (1993). Acquisition of gender-like noun subclasses in an artificial language: The contribution of phonological markers to learning. *Journal of Memory and Language, 32,* 79–95.

Brown, A. S. (1991). A review of the tip-of-the-tongue experience. *Psychological Bulletin, 109,* 204–223.

Brown, R., & McNeill, D. (1966). The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior, 5,* 325–337.

Bruno, N., & Cutting, J. E. (1988). Minimodularity and the perception of layout. *Journal of Experimental Psychology: General, 117,* 161–170.

Buckingham, H. W. (1979). Linguistic aspects of lexical retrieval disturbances in the posterior fluent aphasics. In H. Whitaker & H. A. Whitaker (Eds.), *Studies in neurolinguistics* (Vol. 4, pp. 269–400). New York: Academic Press.

Butterworth, B. (1989). Lexical access and representation in speech production. In W. Masler-Wilson (Ed.), *Lexical representation and process* (pp. 108–135). Cambridge, MA: MIT Press.

Butterworth, B. (1992). Disorders of phonological encoding. *Cognition, 42,* 261–286.

Cacciari, C., Carreiras, M., & Barbolini-Cionini, C. (1997). Anaphora resolution in Italian. *Journal of Memory and Language, 37,* 517–532.

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology, 14,* 177–208.

Caramazza, A., & Miozzo, M. (1997). The relation between syntactic and phonological knowledge in lexical access: Evidence from the "tip-of-the-tongue" phenomenon. *Cognition, 64,* 309–343.

Comrie, B. S. (1978). *The classification of grammatical categories.* Edmonton, Alberta, Canada: Linguistic Research.

Corbett, G. G. (1983). *Hierarchies, target and controllers: Agreement patterns in Slavic.* University Park: Pennsylvania State University Press.

Crain, S., & Steedman, M. (1985). On not being led up the garden path: The use of context by the psychological syntax processor. In D. Dowty, L. Kartunnen, & A. Zwicky (Eds.), *Natural language parsing* (pp. 320–358). Cambridge, England: Cambridge University Press.

Cutler, A. (1982). The reliability of speech error data. In A. Cutler (Ed.), *Slips of the tongue and language production* (pp. 7–28). Amsterdam: Mouton.

Cutler, A., & Norris, D. (1999). Sharpening Ockham's razor. *Behavioral and Brain Sciences, 22,* 40–41.

Cutting, J. C., & Ferreira, V. S. (1999). Semantic and phonological information flow in the production lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 318–344.

Damian, M. F., & Martin, R. M. (1999). Semantic and phonological codes interact in single word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 345–361.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93,* 283–321.

Dell, G. S. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes, 5,* 315–349.

Dell, G. S., Burger, L. K., & Svec, W. R. (1997). Language production and serial order: A functional analysis and a model. *Psychological Review, 104,* 123–147.

Dell, G. S., Reed, K. D., Adams, D. R., & Meyer, A. S. (2000). Speech errors, phonotactic constraints, and implicit learning: A study of the role of experience in language production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1355–1367.

Dell, G. S., & Reich, P. A. (1981). Stages in sentence production: An analysis of speech error data. *Journal of Verbal Learning and Verbal Behavior, 20,* 611–629.

Dell, G. S., Schwartz, M., Martin, M., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review, 104,* 801–838.

Del Viso, S., Igoa, J. M., & García-Albea, J. E. (1987). *Corpus of spontaneous slips of the tongue in Spanish.* Unpublished manuscript.

Del Viso, S., Igoa, J. M., & García-Albea, J. E. (1991). Autonomy of phonological encoding: Evidence from slips of the tongue in Spanish. *Journal of Psycholinguistic Research, 20,* 161–185.

Eberhard, K. M. (1999). The accessibility of conceptual number to processes of subject–verb agreement in English. *Journal of Memory and Language, 41,* 560–578.

Ferreira, F. (1993). Creation of prosody during sentence production. *Psychological Review, 100,* 233–253.

Ferreira, V. S. (1996). Is it better to give than to donate? Syntactic flexibility in language production. *Journal of Memory and Language, 35,* 724–755.

Ferreira, V. S., & Humpreys, K. R. (2001). Syntactic influences on lexical and morphological processing in language production. *Journal of Memory and Language, 44,* 52–80.

Finkel, L. H., & Edelman, G. M. (1989). The integration of distributed cortical systems by reentry: A computer simulation of interactive functionally segregated visual areas. *Journal of Neuroscience, 9,* 3188–3208.

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

Fodor, J. A. (1985). Precis of the modularity of mind. *Behavioral and Brain Sciences, 8,* 1–42.

Foygel, D., & Dell, G. S. (2000). Models of impaired lexical access in speech production. *Journal of Memory and Language, 43,* 182–216.

Franck, J., Vigliocco, G., & Nicol, J. L. (in press). The role of syntactic tree structure and complexity in subject–verb agreement. *Language and Cognitive Processes.*

Frazier, L. (1987). Syntactic processing: A tutorial review. *Attention and Performance, 12,* 559–586.

Frigo, L., & McDonald, J. L. (1998). Properties of phonological markers that affect the acquisition of gender-like subclasses. *Journal of Memory and Language, 39,* 218–245.

Fromkin, V. A. (1971). The non-anomalous nature of anomalous utterances. *Language, 47,* 27–52.

Garnham, A., Shillcock, R. C., Brown, G. D. A., Mill, A. I. D., & Cutler, A. (1981). Slips of the tongue in the London–Lund corpus of spontaneous conversation. *Linguistics, 19,* 805–817.

Garrett, M. F. (1975). The analysis of sentence production. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 133–177). New York: Academic Press.

Garrett, M. F. (1976). Syntactic processes in sentence production. In R. J. Wales & E. C. T. Walker (Eds.), *New approaches to language mechanisms* (pp. 231–256). Amsterdam: North-Holland.

Garrett, M. F. (1980). Levels of processing in sentence production. In B.

Butterworth (Ed.), *Language production: Vol. 1. Speech and talk* (pp. 177–220). London: Academic Press.

Garrett, M. F. (1982). Production of speech: Observations from normal and pathological language use. In A. W. Ellis (Ed.), *Normality and pathology in cognitive functions* (pp. 19–76). London: Academic Press.

Garrett, M. F. (1984). The organization of processing structure for language production: Application to aphasic speech. In D. Caplan, A. R. Lecours, & A. Smith (Eds.), *Biological perspectives on language* (pp. 172–193). Cambridge, MA: MIT Press.

Garrett, M. F. (1993). Errors and their relevance for models of language production. In G. Blanken, J. Dittman, H. Grim, J. Marshall, & C. Wallesch (Eds.), *Linguistic disorders and pathologies* (pp. 72–92). Berlin, Germany: Walter de Gruyter.

Garrett, M. F. (2000). Remarks on the architecture of language processing systems. In Y. Grodzinsky, L. Shapiro, & D. Swinney (Eds.), *Language and brain* (pp. 31–69). San Diego, CA: Academic Press.

Glaser, W. R., & Düngelhoff, F.-J. (1984). The time course of picture–word interference. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 640–654.

Goodglass, H., Kaplan, E., Weintraub, S., & Ackerman, N. (1976). The "tip of the tongue" phenomenon in aphasia. *Cortex, 12,* 145–153.

Griffin, Z. M., & Bock, K. (1998). Constraint, word frequency, and the relationship between lexical processing levels in spoken word production. *Journal of Memory and Language, 38,* 313–338.

Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science, 11,* 274–279.

Grossberg, S. (1976). Adaptive pattern classification and universal recoding: II. Feedback, expectation, olfaction, and illusions. *Biological Cybernetics, 23,* 187–202.

Grossberg, S. (2001). The complementary brain: A unifying view of brain specialization and modularity. *Trends in Cognitive Sciences, 4,* 233–246.

Harley, T. A. (1984). A critique of top-down independent level models of speech production: Evidence from non-plan-internal speech errors. *Cognitive Speech, 8,* 191–219.

Harley, T. A. (1993). Phonological activation of semantic competitors during lexical access in speech production. *Language and Cognitive Processes, 8,* 291–309.

Harley, T. A., & Bown, H. E. (1998). What causes a tip-of-the-tongue state? Evidence for local neighbourhood effects in speech production. *British Journal of Psychology, 89,* 151–174.

Harley, T. A., & MacAndrew, S. B. G. (1995). Interactive models of lexicalization: Some constraints from speech error, picture naming, and neuropsychological data. In J. Levy, D. Bairaktaris, J. Bullinara, & D. Cairns (Eds.), *Connectionist models of memory and language* (pp. 311–331). London: UCL Press.

Hartsuiker, R. J., Antón-Méndez, I., & Van Zee, M. (2001). Object attraction in subject–verb agreement construction. *Journal of Memory and Language, 45,* 546–572.

Hartsuiker, R. J., & Kolk, H. H. J. (1998a). Syntactic facilitation in agrammatic sentence production. *Brain and Language, 62,* 221–254.

Hartsuiker, R. J., & Kolk, H. H. J. (1998b). Syntactic persistence in Dutch. *Language & Speech, 41,* 143–184.

Hartsuiker, R. J., & Kolk, H. H. J. (2001). Error monitoring in speech production: A computational test of the perceptual loop theory. *Cognitive Psychology, 42,* 113–157.

Hartsuiker, R. J., Kolk, H. H. J., & Huinck, W. J. (1999). Subject–verb agreement construction in agrammatic aphasia: The role of conceptual number. *Brain and Language, 69,* 119–160.

Hartsuiker, R. J., Kolk, H. H. J., & Huiskamp, P. (1999). Priming word order in sentence production. *Quarterly Journal of Experimental Psychology, 52A,* 129–147.

Hartsuiker, R. J., Schriefers, H. J., Bock, J. K., & Kikstra, G. M. (2000). *Morphophonological influences on the construction of subject–verb agreement.* Manuscript submitted for publication.

Hartsuiker, R. J., & Westenberg, C. (2000). Word order priming in written and spoken sentence production. *Cognition, 75,* B27–B39.

Henaff Gonon, M., Bruckert, R., & Michel, F. (1989). Lexicalization in an anomic patient. *Neuropsychologica, 27,* 391–407.

Howard, D., & Orchard-Lisle, V. (1984). On the origin of semantic errors in naming: Evidence from the case of a global aphasic. *Cognitive Neuropsychology, 1,* 163–190.

Iwasaki, N., Vigliocco, G., & Garrett, M. F. (1998). Adjectives and adjectival nouns in Japanese: Psychological processes in sentence production. In D. J. Silva (Ed.), *Japanese/Korean Linguistics* (Vol. 8, pp. 93–106). Stanford, CA: Center for the Study of Language and Information.

Jakobson, R. (1959). Boas' view of grammatical meaning. *American Anthropologist, 61,* 139–145.

Jescheniak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 824–843.

Jescheniak, J. D., & Schriefers, H. (1998). Discrete serial versus cascaded processing in lexical access in speech production: Further evidence from the coactivation of near synonyms. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 1256–1274.

Johnson-Laird, P. N. (1983). *Mental models.* Cambridge, England: Cambridge University Press.

Kay, J., & Ellis, A. (1987). A cognitive neuropsychological study of anomia. *Brain, 110,* 613–629.

Keenan, E. L., & Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry, 8,* 63–99.

Kelly, M. H. (1992). Using sound to solve syntactic problems: The role of phonology in grammatical category assignments. *Psychological Review, 99,* 349–364.

Kempen, G. (1999). *Human grammatical coding.* Manuscript in preparation.

Kempen, G., & Hoenkamp, E. (1987). Incremental procedural grammar for sentence formulation. *Cognitive Science, 11,* 201–258.

Kempen, G., & Huijbers, P. (1983). The lexicalization process in sentence production and naming: Indirect election of words. *Cognition, 14,* 185–209.

Kempen, G., & Vosse, T. (1989). Incremental syntactic tree formation in human sentence processing: A cognitive architecture based on activation decay and simulated annealing. *Connection Science, 1,* 273–290.

Kuminiak, F., & Badecker, W. (1998, November). *Production of subject–verb agreement and the role of case marking.* Paper presented at the meeting of the Psychonomic Society, Dallas, TX.

Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M. (1995). Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision Research, 35,* 389–412.

Laver, J. (1980). Monitoring systems in the neurolinguistic control of speech production. In V. A. Fromkin (Ed.), *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand* (pp. 287–305). New York: Academic Press.

Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition, 14,* 41–104.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation.* Cambridge, MA: MIT Press.

Levelt, W. J. M., & Maassen, R. (1981). Lexical search and order of mention in sentence production. In W. Klein & W. Levelt (Eds.), *Crossing the boundaries in linguistics* (pp. 221–252). Dordrecht, the Netherlands: Reidel.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22,* 1–75.

Levelt, W. J. M., Schriefers, H., Vorberg, D., Meyer, A. S., Pechmann, T., & Havinga, J. (1991). The time course of lexical access in speech

production: A study of picture naming. *Psychological Review, 98,* 122–142.

Lloyd-Jones, T. J., & Humphreys, W. G. (1997). Perceptual differentiation as a source of category effects in object processing: Evidence from naming and object decision. *Memory & Cognition, 25,* 18–35.

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

MacKay, D. G. (1992). Awareness and error detection: New theories and research paradigms. *Consciousness and Cognition, 1,* 199–225.

Martin, N., Gagnon, D. A., Schwartz, M. F., Dell, G. S., & Saffran, E. M. (1996). Phonological facilitation of semantic errors in normal and aphasic speakers. *Language and Cognitive Processes, 11,* 257–282.

Martin, R. C., Vigliocco, G., & Garrett, M. F. (1999). Access to count/mass information without access to phonology in an anomic patient. *Brain and Language, 46,* 426–428.

Martin, N., Weisberg, R. W., & Saffran, E. M. (1989). Variables influencing the occurrence of naming errors: Implications for models of lexical retrieval. *Journal of Memory and Language, 28,* 462–485.

Marx, E. (1999). Gender processing in speech production: Evidence from German speech errors. *Journal of Psycholinguistic Research, 28,* 601–621.

Mattson, M. E., & Baars, B. J. (1992). Error-minimizing mechanisms: Boosting or editing? In B. J. Baars (Ed.), *Experimental slips and human error: Exploring the architecture of volition* (pp. 263–287). New York: Plenum.

Meyer, A. S., & Bock, K. (1992). The tip-of-the-tongue phenomenon: Blocking or partial activation? *Memory & Cognition, 20,* 715–726.

Miceli, G., & Caramazza, A. (1988). Dissociation of inflectional and derivational morphology. *Brain & Language, 35,* 24–65.

Miozzo, M., & Caramazza, A. (1997). The retrieval of lexical-syntactic features. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 1410–1423.

Nakayama, K., & Shimojo, S. (1992, September 4). Experiencing and perceiving visual surfaces. *Science, 257,* 1357–1363.

Newell, R. S. (1973). You can't play 20 questions with nature and win: Projective comments on the papers of this symposium. In W. G. Chase (Ed.), *Visual information processing* (pp. 283–308). New York: Academic Press.

Nooteboom, S. G. (in press). Listening to oneself: Monitoring speech production. In R. J. Hartsuiker, R. Bastiaanse, A. Postma, & F. N. K. Wijnen (Eds.), *Phonological encoding and self-monitoring in normal and pathological speech.* Hove, England: Psychology Press.

Norris, D., McQueen, J., & Cutler, A. (2000). Merging information in speech: Feedback is never necessary. *Behavioral and Brain Sciences, 23,* 299–370.

Peterson, R. R., & Savoy, P. (1998). Lexical selection and phonological encoding during language production: Evidence for cascaded processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 539–557.

Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language, 39,* 633–651.

Postma, A. (2000). Detection of errors during speech production: A review of speech monitoring models. *Cognition, 77,* 97–131.

Postma, A., & Kolk, H. H. J. (1993). The covert repair hypothesis: Prearticulatory repair processes in normal and stuttered disfluencies. *Journal of Speech and Hearing Research, 36,* 472–487.

Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review, 107,* 460–499.

Roelofs, A., Meyer, A. S., & Levelt, W. J. M. (1996). Interaction between semantic and orthographic factors in conceptually driven naming: Comment on Starreveld and La Heij (1995). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 246–251.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996, December 13). Statistical learning by 8-month-old infants. *Science, 274,* 1926–1928.

Schade, U., & Laubenstein, U. (1993). Repairs in a connectionist language-production model. In R. Köhler & B. B. Rieger (Eds.), *Contributions to quantitative linguistics* (pp. 79–90). Dordrecht, the Netherlands: Kluwer Academic.

Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: Picture–word interference studies. *Journal of Memory and Language, 29,* 86–102.

Seidenberg, M., & McDonald, M. C. (1999). A probabilistic constraints approach to language acquisition and processing. *Cognitive Science, 23,* 569–588.

Shillcock, R. (2000, May). *Spoken word access: Evidence from statistical analyses of the lexicon.* Paper presented at the Spoken Word Access Processing workshop, Nijmegen, the Netherlands.

Silverberg, N., Vigliocco, G., Insalaco, D., & Garrett, M. F. (1998). When reading a sentence is easier than reading a "little" word: The role of production processes in deep dyslexia. *Aphasiology, 12,* 335–356.

Slobin, D. I. (1992). *The cross-linguistic study of language acquisition* (Vol. 3). Hillsdale, NJ: Erlbaum.

Slobin, D. I. (1996). From "thought and language" to "thinking for speaking." In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70–96). New York: Cambridge University Press.

Spitzer, M. (1999). *The mind in the net.* Cambridge, MA: MIT Press.

Starreveld, P. A., & La Heij, W. (1995). Semantic interference, orthographic facilitation, and their interaction in naming tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 686–698.

Starreveld, P. A., & La Heij, W. (1996). Time-course analysis of semantic and orthographic context effects in picture naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 896–918.

Starreveld, P. V., & La Heij, W. (1999). What about phonological facilitation, response-set membership, and phonological coactivation? *Behavioral and Brain Sciences, 22,* 56–57.

Stemberger, J. P. (1982). *The lexicon in a model of language production.* Unpublished doctoral dissertation, University of California, San Diego.

Stemberger, J. P. (1985). An interactive activation model of language production. In A. W. Ellis (Ed.), *Progress in the psychology of language* (Vol. 1, pp. 143–186). London: Erlbaum.

Tabor, W., & Tanenhaus, M. (1999). Dynamical sentence processing. *Cognitive Science, 23,* 491–515.

Taft, M., & Meunier, F. (1995, November). *A challenge for neural networks: Gender processing in French.* Paper presented at the meeting of the Psychonomic Society, Los Angeles.

Tanenhaus, M. K., & Trueswell, J. C. (1995). Sentence comprehension. In J. Miller & P. Eimas (Eds.), *Handbook of perception and cognition: Speech, language and cognition* (pp. 217–262). San Diego, CA: Academic Press.

Teschner, R. V. (1987). The genders of French and Spanish noun cognates—Some statistics and a list. *Canadian Modern Language Review/Revue Canadienne des Langues Vivantes, 43,* 256–266.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language, 33,* 285–318.

Tucker, G. R., Lambert, W. E., & Rigault, A. A. (1977). *The French speaker's skill with grammatical gender: An example of rule-governed behavior.* The Hague, the Netherlands: Mouton.

Van Wijk, C., & Kempen, G. (1987). A dual system for producing self-repairs in spontaneous speech: Evidence from experimentally elicited corrections. *Cognitive Psychology, 19,* 403–440.

Vigliocco, G. (1995). *Quando il soggetto e il verbo sono in disaccordo: Uno studio sperimentale sui processi di codifica grammaticale in produzione* [When the subject and verb disagree: An experimental study of the

processes involved in grammatical encoding during production]. Unpublished doctoral dissertation, University of Trieste, Trieste, Italy.

Vigliocco, G. (2000). *A corpus of slips of the tongue in Italian.* Unpublished data.

Vigliocco, G., Antonini, T., & Garrett, M. F. (1997). Grammatical gender is on the tip of Italian tongues. *Psychological Science, 8,* 314–317.

Vigliocco, G., Butterworth, B., & Garrett, M. F. (1996). Subject–verb agreement in Spanish and English: Differences in the role of conceptual constraints. *Cognition, 61,* 261–298.

Vigliocco, G., Butterworth, B., & Semenza, C. (1995). Constructing subject–verb agreement in speech: The role of semantic and morphological factors. *Journal of Memory and Language, 34,* 186–215.

Vigliocco, G., & Franck, J. (1999). When sex and syntax go hand in hand: Gender agreement in language production. *Journal of Memory and Language, 40,* 455–478.

Vigliocco, G., & Franck, J. (2001). When sex affects syntax: Context effects in sentence production. *Journal of Memory and Language, 45,* 368–390.

Vigliocco, G., Franck, J., Antón-Méndez, I., & Collina, S. (2001). *The interplay between syntax and form in language production: An investigation of gender agreement in Romance languages.* Manuscript submitted for publication.

Vigliocco, G., Hartsuiker, R. J., Jarema, G., & Kolk, H. H. J. (1996). One or more labels on the bottles? Notional concord in Dutch and French. *Language and Cognitive Processes, 11,* 407–442.

Vigliocco, G., Lauer, M., Damian, M., & Levelt, W. (2002). Semantic and syntactic forces in noun phrase production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 46–58.

Vigliocco, G., & Nicol, J. (1998). Separating hierarchical relations and word order in language production: Is proximity concord syntactic or linear? *Cognition, 68,* 13–29.

Vigliocco, G., Vinson, D. P., Indefrey, P., & Levelt, W. J. M. (2000, November). *The interplay of meaning and syntax in speech: Insights from speech errors.* Paper presented at the meeting of the Psychonomic Society, New Orleans, LA.

Vigliocco, G., Vinson, D., Martin, R., & Garrett, M. F. (1999). Is "count" and "mass" information available when the noun is not? An investigation of tip of the tongue states and anomia. *Journal of Memory and Language, 40,* 534–558.

Vigliocco, G., & Zilli, T. (1999). Syntactic accuracy in sentence production: Gender disagreement in Italian language impaired and unimpaired speakers. *Journal of Psycholinguistic Research, 28,* 623–648.

Vigliocco, G., & Zorzi, M. (1999). Contact points between lexical retrieval and sentence production. Commentary to "A theory of lexical access in speech production" by Levelt, Roelofs and Meyer. *Behavioral and Brain Sciences, 22,* 58–59.

Vinson, D. P., & Vigliocco, G. (1999). Can independence be observed in a dependent system? The case of tip-of-the-tongue states. *Brain and Language, 68,* 118–126.

Vinson, D. P., & Vigliocco, G. (2002). A semantic analysis of grammatical class impairments. *Journal of Neurolinguistics, 15,* 317–351.

Vosse, T., & Kempen, G. (2000). Syntactic structure assembly in human parsing: A computational model based on competitive inhibition and a lexicalist grammar. *Cognition, 75,* 105–143.

Wheeldon, L. R., & Levelt, W. J. M. (1995). Monitoring the time course of phonological encoding. *Journal of Memory and Language, 34,* 311–334.

Zorzi, M., & Vigliocco, G. (1999). Compositional semantics and the lemma dilemma. Commentary to "A Theory of Lexical Access in Speech Production" by Levelt, Roelofs and Meyer. *Behavioral and Brain Sciences, 22,* 60–61.