The process of spoken word recognition: An introduction

LORRAINE KOMISARJEVSKY TYLER
University of Cambridge

ULI H. FRAUENFELDER*
Max-Planck-Institut für Psycholinguistik, Nijmegen

Abstract

This introduction sets the stage for the papers making up this special issue. Its focus is on two major problems in the study of lexical processing—determining the phases involved in recognising a spoken word and identifying the nature of different types of contextual influences on these phases. An attempt is made to decompose the process of recognising a word into phases which have both theoretical and empirical consequences. A similar analytic approach is taken in the discussion of the problem of context effects by distinguishing qualitatively different types of context (lexical, intra-lexical, syntactic, semantic, and interpretative). We argue that such an approach is necessary to make explicit the relationship between a particular type of contextual information and the phase(s) of processing at which it has its impact.

1. Introduction

Until recently, *spoken* word recognition has been a neglected area of study. Psychological research has focused on the *written* word, and generally has assumed that the processes involved in this modality were also those involved in the auditory modality. Thus, early models of word recognition (e.g., Becker & Killion, 1977; Forster, 1976; Morton, 1969) were developed on the basis of data obtained in reading tasks, although they were assumed to pro-

*Both authors contributed equally to this Introduction and to the Special Issue. Reprint requests should be sent to Uli Frauenfelder, Max-Planck-Institut für Psycholinguistik, Postbus 310, 6500AH Nijmegen, The Netherlands.

vide a general description of word recognition processes which was not modality-specific. More recently, there has been increased emphasis on developing models which are intended to account specifically for data on spoken language processing (e.g., Cole & Jakimik, 1980; Elman & McClelland, 1984; Marslen-Wilson, 1987, this issue; Marslen-Wilson & Tyler, 1980; Marslen-Wilson & Welsh, 1978) and to evaluate the relationship between processing in the two modalities (e.g., Bradley & Forster, 1987, this issue).

Apart from these developments in psycholinguistics, there have also been advances in speech perception, automatic speech recognition, linguistic theory and parsing that are all relevant to the problem of auditory word recognition.

In the area of speech perception, for example, there is growing awareness that the processing of the acoustic-phonetic input must be studied within the context of the lexical processing system, rather than merely within the phonetic domain (Nooteboom, 1979; Pisoni & Luce, 1987, this issue). In the field of automatic speech recognition, attempts to represent phonetic and phonological information formally and to develop algorithms for using this information in the analysis of the acoustic signal (e.g., Church, 1987, this issue) complement the efforts by psycholinguists to make mental structure and process explicit in models of lexical processing.

To capture structural regularities at different linguistic levels, phonological theory is moving away from the use of linear representations towards more hierarchical ones. Psycholinguists are taking increasing interest in such representations and are beginning to explore the possibility that listeners use phonological and prosodic knowledge to parse the sensory input during word recognition (cf. Frazier, 1987, this issue; Grosjean & Gee, 1987, this issue). The study of morphology has resulted in new theories of the structure and organisation of lexical entries that have provided important hypotheses for psycholinguistic research (e.g., Aronoff, 1976; Bybee, 1985; Selkirk, 1984), as has the development of grammatical theories which attribute considerable structural information to the lexicon (e.g., Bresnan, 1978). Finally, recent research in parsing has focused on specifying the informational content of lexical entries necessary to develop parsers which can use lexical representations for a word-by-word construction of higher-level representations (e.g., Ford, Bresnan & Kaplan, 1982).

Such a multi-disciplinary approach is necessary to understand the lexical processing system and the way in which it relates sound to meaning. Although not all of the work in these various disciplines is explicitly formulated as investigating psychological issues, it does, nevertheless, represent a body of knowledge which is invaluable in the development of psychological models of word recognition.

2. Lexical processing

In the following sections we will first outline what we consider to be the major phases involved in lexical processing and indicate how different theoretical positions have dealt with each of them. Then we will present a brief overview of the way in which context effects of different types have been assumed to intervene in these phases of lexical processing. Throughout this introduction we will raise some of the issues that continue to dominate research in lexical representation and process.

One of our objectives is to confront the terminological confusion plaguing word recognition research. Indeed, it is clear, even from the contributions to this volume, that we are still far from a terminological consensus. Basic terms like "word recognition" and "lexical access" are often used to refer to very different processes (cf. Tanenhaus & Lucas, 1987, this issue). We will attempt to identify in as theoretically neutral a fashion as possible the major aspects of lexical processing in an effort to develop a terminology which is consistent with various theoretical frameworks.

2.1. Initial lexical contact

The process of recognising a spoken word begins when the sensory input—or, more precisely, some representation computed from this input—makes initial contact with the lexicon. In this *initial contact phase*, the listener takes the speech wave as input and generates the representation(s) which contact the internally stored form-based representations associated with each lexical entry. A major question concerns the nature of the representation which makes contact with the lexicon. This representation has important consequences not only for *which* lexical entries are initially contacted but also for *when* they are contacted.

2.1.1. Contact representations

Many different *contact representations* have been proposed to mediate this initial phase—ranging from temporally defined spectral templates (e.g., Klatt, 1980) to abstract linguistic units like phonemes (e.g., Pisoni & Luce, 1987, this issue) or syllables (e.g., Mehler, 1981). The properties of these representations have potential consequences for the size of the initially contacted subset of the lexicon. The richer or more discriminative the information in the contact representation, the smaller the number of lexical entries initially contacted. To illustrate this point, we can contrast the effect of a phoneme-based representation with that of a robust feature representation (cf. Shipman & Zue, 1982) in which only six classes of phonemes are distinguished. In the former case, because the description of the input is much richer, it is more discriminative, and the size of the initially contacted set is smaller than in the latter case.

2.1.2. When initial contact occurs

The amount of speech required to compute the contact representation determines the moment at which initial contact can occur. Clearly, the longer the stretch of speech signal that the system needs to accumulate to construct this representation, the more the initial contact is delayed. We can contrast models with potentially immediate contact such as the LAFS model (Klatt, 1980) in which the first 10 ms spectral template initiates a path to a lexical entry, with models in which there is a much longer "dead period" during which no contact is possible. Consistent with the latter type of proposal, it has been suggested that the first syllable of a word (Bradley & Forster, 1987, this issue) or the first 150 ms of a word (Marslen-Wilson, 1984; Salasoo & Pisoni, 1985; Tyler, 1984) needs to be analysed before contact can be made.

In some models the first contact with lexical entries is based upon some initial portion of a word (Cole & Jakimik, 1980, Marslen-Wilson & Welsh, 1978). In the "cohort model", for example, the "word-initial cohort" contains all of the words in a language matching some beginning portion of the input (Marslen-Wilson & Tyler, 1980; Marslen-Wilson & Welsh, 1978). This view, in which priority is given to *temporally early* information, can be contrasted with approaches in which information which is physically *more salient*—irrespective of its temporal location—is used to contact the lexicon. For instance, Grosjean and Gee (1987, this issue) claim that stretches of the signal that are particularly reliable (such as stressed syllables) establish the initially contacted subset of the lexicon. These approaches all share the assumption that there is a discrete stage of initial contact which delimits a subset of the lexicon.

2.1.3. Advantages and disadvantages of discrete initial contact

The obvious advantage of discrete initial contact is that not all the entries in the lexicon need to be considered in subsequent phases of analysis. However, there are problems associated with the assumption that lexical entries are only ever considered if they are included in the initial subset of words matching the contact representation. For example, the intended word will never be located when the contact representation is misperceived. In order to reduce the risk of such unsuccesful initial contact, the contact representation has to be constrained. It has to be broad enough to ensure that the intended word is contacted, and yet specific enough so that only a minimal number of entries is contacted.

The contact representation must also be segmented correctly; it should correspond exactly to that portion of the stored lexical representation with which it is to be matched. If, for instance, a stretch of speech not corresponding to the initial part of a word is used to make contact with the beginnings of stored lexical representations, an inappropriate subset of the lexicon will be contacted, and the intended word will not be recognised (Frauenfelder, 1985). Positional or segmentation information, either in the contact representation itself or in the preceding context, must be available to ensure that proper initial contact takes place.

Models which do not assume a unique contact for each word avoid these potential problems. For example, the Trace model (Elman & McClelland, 1984) allows each activated phoneme to define a new subset of lexical entries containing this phoneme. The set of activated lexical entries constantly changes as old members drop out and new members are added. This avoids the problem of excluding the intended word from the pool of activated candidates, although it runs the risk of having too many activated words at each moment in time, making it more difficult to narrow in on and select the correct word. It remains an important—and unresolved—question whether or not word recognition does take the form of narrowing-down process of an initially established subset of the lexicon.

2.2. Activation

The lexical entries that match the contact representation to some criterial degree during the initial contact phase are assumed to change in state. In the absence of a theoretically neutral term for this change, we will refer to it as "activation". Theories differ in the claims they make concerning the factors that determine the relative status of activated words. For instance, the original version of the cohort theory proposed that all lexical entries matching the contact representation were equally activated and therefore had equal status. In the search model described by Bradley and Forster (1987, this issue), the relative status (the term "level of activation" is not appropriate for this model) of lexical entries at lexical contact depends upon properties of these entries themselves—in particular, upon their frequency of occurrence in the language. Lexical entries are ordered (within their respective subset or "bins") according to frequency. In other models, such as the current version of the cohort theory (Marslen-Wilson, 1987, this issue) and the Trace model, the degree of activation of a contacted lexical entry varies depending on both its goodness of fit with the contact representation(s) and its own internal specifications (e.g., frequency of occurrence).

2.3. Selection

After initial contact and activation of a subset of the lexicon, accumulating sensory input continues to map onto this subset until the intended lexical entry is eventually selected. This *selection phase* has been described in various ways: As a process of *differentiation* (McClelland & Rumelhart, 1986), *reduction* (Marslen-Wilson, 1984) or *search* (Forster, 1976). In the Trace model, the differential activation of lexical entries provides the basis for selection. Through processes of activation and inhibition, one entry eventually emerges as the most activated relative to all other entries. In contrast to this approach, the original formulation of the cohort theory saw this as an all-or-none process. The internal specifications of lexical entries were assessed against the sensory input and those which failed to match dropped out of the cohort. Thus, entries were either in or out of the cohort. A rather different approach is taken in the most recent version of the cohort theory (Marslen-Wilson, 1987, this issue) where lexical entries failing to match the input are not dropped from the cohort completely, but rather their level of activation starts to decay in the absence of further bottom-up support. In search models, the correct word is selected by a process which searches through the frequency-ordered set of lexical entries (Bradley & Forster, 1987, this issue).

2.4. Word recognition

We will reserve the term *word recognition* for the end-point of the selection phase when a listener has determined which lexical entry was actually heard. An important objective in approaches which emphasise the temporal nature of the recognition process, has been to determine the *word recognition point*, that is, the precise moment in time at which a word is recognised.

It is widely accepted that listeners generally recognize words, either in isolation or in context, before having heard them completely (Grosjean, 1980; Marslen-Wilson, 1984; Marslen-Wilson & Tyler, 1980). The exact recognition point of any given word depends upon a number of factors including its physical properties (e.g., length, stimulus quality), its intrinsic properties (frequency), the number and nature of other words in the lexicon that are physically similar to this word (i.e., its competitors or fellow cohort members) and the efficiency of the selection process. If the simplifying assumption is made that the acoustic signal is recognised sequentially, categorically and correctly as a sequence of discrete segments (e.g., phonemes or syllables) and that the selection process retains only those lexical entries matching this sequence, then it is possible to determine the recognition point for each word. In this case, a word's recognition point corresponds to its *uniqueness point*—

that is, the point at which a word's initial sequence of segments is common to that word and no other. If, however, the analysis of the input proceeds in a probabilistic rather than categorical fashion, then a word is not necessarily recognised at the uniqueness point, but rather later at the moment the sensory input matches one single lexical candidate better than all others by some criterial amount (Marcus & Frauenfelder, 1985).

2.5. Lexical access

The goal of lexical processing is to make available the stored knowledge associated with a word (cf. Johnson-Laird, 1987, this issue) so that this can be used to develop a meaningful interpretation of an utterance. We use the term *lexical access* to refer to the point at which the various properties of stored lexical representations—phonological, syntactic, semantic, pragmatic—become available. One central question is *when* does this lexical information become available to the rest of the language processing system?

Most theories agree that some form-based information must be available in the initial contact phase of lexical processing—otherwise there would be no basis for a match with the sensory input. There is disagreement, however, on the point at which other types of stored lexical knowledge become available. The range of different views is exemplified by the contrast between the cohort and search models. In the cohort model, all stored information is activated simultaneously upon initial contact (Marslen-Wilson & Tyler, 1980). In the search model, although some form-based description must be made available early in the process (upon initial contact), stored syntactic and semantic information does not become available until a word is accessed and recognised (Forster, 1976, 1979). This is because such information is stored centrally in a master file which is not entered until the word has been recognised (a process which takes place in the access "bin").

The assumed relationship between lexical access and word recognition varies depending upon the theory. In models like that of Bradley and Forster (1987, this issue), lexical access and word recognition, as defined here, are indistinguishable (although the authors themselves introduce another theoretical distinction between the two) since lexical information becomes available (lexical access) only when a single lexical entry has been found (word recognition). In models like the cohort model, there is a clear difference in that lexical access precedes word recognition.

Up to now we have only discussed the phases involved in recognising words and accessing stored lexical information. What remains to be considered now is how higher-order context influences spoken word recognition.

3. Context effects

An ubiquitous finding in the literature is that context plays an important role in spoken word recognition (e.g., Blank & Foss, 1978; Foss, 1982; Cairns & Hsu, 1980; Marslen-Wilson & Tyler, 1980, Salasoo & Pisoni, 1985). To explain the ease and rapidity with which listeners recognise words, psycholinguistics often appeal to context. Their general claim is that lexical processing depends on two broad classes of information—representations computed from the sensory input, and those constructed from the previous context using higher sources of knowledge (e.g., lexical, syntactic, semantic, and pragmatic). To understand the nature of these contextual influences we need to specify the answers to at least three related questions: (1) *which* types of context affect lexical processing? (2) *when* do these contexts influence specific processes involved in recognising a word? and (3) *how* do these types of context have their effects? The answers to these questions have important implications for the general structure of the language processing system.

3.1. Autonomy versus interaction

There are currently two strongly opposing views concerning the structure of the language processing system—the autonomous and the interactive views. Each provides different answers to questions about the way in which information can flow through the language processing system. According to autonomous theories, there are strong constraints on the way in which contextual information can affect the bottom-up analysis. Context *cannot* have its effect prior to the completion of the phases of lexical processing leading up to word recognition. It only contributes to the evaluation and integration of the output of lexical processing, but not to the generation of this output (e.g., Forster, 1979; Norris, 1986; Tanenhaus, Carlson & Seidenberg, 1984; Tanenhaus & Lucas, 1987, this issue). However, such models permit "lateral" flow of information *within* a given processing level (e.g., between words within the lexicon).

Interactive models, in contrast, allow different kinds of information to interact with each other. However, the extent to which contextual information is allowed to intervene in any of the phases of lexical processing varies considerably in different theories. In certain interactive models (e.g., Morton, 1969), expectations generated from higher-level representations actually intervene directly in the earliest phases of lexical processing by altering the activation of lexical elements. In others, context only operates on a subset of elements selected on the basis of the sensory input (e.g., Marslen-Wilson & Welsh, 1978).

At first glance, the predictions of these two classes of theories appear to be clear-cut, and choosing between them straightforward. In reality, however, differentiating between the models is extremely difficult. A major problem is that the distinction between autonomous and interactive models is not dichotomous but continuous. Models are autonomous or interactive to varying degrees. Consequently, there are relatively large differences between models *within* the autonomous or interactive class, and very small differences *between* some autonomous and interactive models (cf. Tyler & Marslen-Wilson, 1982).

For example, autonomous models vary in the extent to which the principle of autonomy constrains the system's operations. This is reflected in differences in the size and number of postulated autonomous modules. In Forster's model (1979), for example, there are several autonomous processing modules, each corresponding to a putative linguistic level (lexical, syntactic, and semantic). These modules are configured serially so that each module only receives and processes the output of the immediately lower level; any information derived from higher processing modules cannot affect processing operations at the lower level. In contrast, there is only a single module (the language processing module) in Fodor's (1983) model. Information can flow freely between the different subcomponents of the language processing module, but information coming from outside the language processor (knowledge of the world), cannot alter the course of its operations. In this model, then, the principle of autonomy applies only to the language processing module as a whole and does not constrain the intermediate levels of analysis, as in Forster's model.

Interactive models are just as varied as autonomous models. They primarily differ as a function of how (or how much) contextual information is permitted to affect the sensory analysis. On interactive accounts, both contextual and sensory information contribute to lexical processing. Context can propose lexical candidates for consideration even before any sensory input has been received (Morton, 1969). Other views, in which the flow of information is more highly constrained, allow context only to dispose of candidates and not to propose them (e.g., Marslen-Wilson, 1987, this issue). Lexical candidates which are contextually appropriate are integrated into the higher-level representation. Those which are contextually inappropriate are either completely eliminated from contention during the selection phase (Marslen-Wilson & Welsh, 1978), or their activation levels gradually decay (Marslen-Wilson, 1987, this issue).

Because there is no standard autonomous or interactive theory, it is often difficult in practice to distinguish between these two classes of models. Results that have been taken to favour interactive models can be explained by autonomous models making slightly different assumptions. For example, certain context effects can be accounted for by an autonomous model in which multiple, rather than single, outputs are passed on to higher-level modules (e.g., Norris, 1982, 1986). These outputs are then evaluated against the context in a parallel fashion. The lexical candidates that best match contextual requirements are recognised even in the absence of sufficient sensory input to uniquely specify the word. The process of word recognition is thus claimed to be unaffected by context, and the autonomy hypothesis remains unscathed. However, as Tyler and Marslen-Wilson (1982) point out, the assumptions of autonomy are so weakened in such a model that it essentially becomes indistinguishable empirically from an interactive model.

Most autonomous models do, however, make predictions about the nature of context effects which are empirically distinguishable from those proposed by interactionist models. The major difference between the models lies in the claims each makes about the *moment* at which context has its effect. For context to exert its influence *before* or *during* the selection phase constitutes evidence against a strong autonomy view and in favour of certain types of interactionist view.

Morton (1969) and Grosjean (1980), among others, have advanced strong interactionist positions by claiming that context can have an effect on lexical processing even before any sensory input is heard. This is achieved in Morton's model by allowing context to increase the level of activation of individual logogens even before the sensory input makes contact with the lexicon. As a result, less sensory input is needed to reach threshold and to recognise a contextually appropriate compared to a contextually inappropriate word. For Grosjean, rather than affecting threshold levels, context narrows down the set of words which are matched against the sensory input to those which are contextually appropriate. It is not made clear in either of these models how context can function in advance of any sensory input. Unless context effects operate via spreading activation, they can help select contextually appropriate lexical entries only if the syntactic and semantic properties of these lexical entries are already available for contextual evaluation. But how can they be when no sensory input corresponding to any lexical entry has yet entered the system? For such a system to work without spreading activation requires the stored semantic and syntactic information of *all* entries in the entire lexicon to always be available to be assessed against the context.

In most other interactionist models, context effects occur at a later phase of the process—when the sensory input has made initial contact with the lexicon. These models claim that this is the earliest moment in time at which context can exert its influence on lexical processing because this is when the stored properties of words first become available. It is on the basis of these stored representations that a word can be evaluated for its contextual appropriateness. So, for example, in the original version of the cohort theory, it was at the point of initial contact that word candidates were first evaluated for their contextual appropriateness. Candidates whose internal specifications were incompatible with the context dropped out of the activated subset. This process facilitated the selection of a single candidate from amongst the initially contacted set by reducing the set to only those words which were contextually appropriate. This resulted in earlier recognition for a word in context than in isolation. Word recognition in this model, then, was seen as involving both autonomous and interactive processes. The initial process of contacting the lexicon was autonomous, but the selection process was interactive.

Autonomy theorists in contrast, claim that it is only *after* a word emerges as the single best fit with the sensory input that context can begin to have an effect. For example, Forster (1976) claims that when the sensory input uniquely specifies a word, the pointer for this word contacts the appropriate entry in the master file. It is in the master file that a word's internal properties are stored and, therefore, it is only at this point that the syntactic and semantic properties of the word are evaluated against the specifications of the context. The role of context, therefore, is restricted to the post-access phase of lexical processing (e.g., Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982; Swinney, 1979).

To evaluate these various theoretical predictions, it is essential to establish the precise moment in processing at which context produces its effect. The recent development of on-line techniques has provided psycholinguists with increased power to resolve the temporal properties of word recognition necessary to determine the loci of context effects. Nonetheless, the use of these experimental techniques introduces new problems that make data interpretation difficult. In the first flush of enthusiasm over these techniques, any on-line task was thought to be as good as any other. The possibility that different tasks might tap different phases of lexical processing was largely ignored. However, individual tasks are now coming under greater scrutiny. This can be seen in the current debate over the appropriate use of naming and lexical decision tasks. The lexical decision task appears to be sensitive to late decision processes (e.g., Forster, 1979; Jakimik, Cole & Rudnicky, 1985; Seidenberg, Waters, Sanders & Langer, 1984) and is not, therefore, appropriate for assessing the role of context in the early phases of lexical processing. Naming, in contrast, seems to reflect earlier phases of lexical processing and therefore promises to be a more useful task with which to determine the locus of context effects.

3.2. Different types of context effects

The proper characterisation of context effects is a delicate task. Nonetheless, since context undeniably does affect lexical processing at some point, it is essential to determine the specific nature of this influence if we are to arrive at a complete understanding of the language processing system. In the following sections, we will distinguish between different types of context and examine the ways in which each has been argued to affect various phases of lexical processing. We will draw a broad distinction between two types of context which we will call *structural* and *non-structural*.

We can define structural context as that which results from constraints on the ways in which elements can be combined into higher-level units. This type of constraint can apply at the phoneme, morpheme, phrase, utterance and discourse levels. So, for example, the rules determining the legal combination of morphemes into polymorphemic words constitute one type of structural constraint. In principle, this is similar to the constraints governing the set of elements which can combine to form, for example, a noun phrase or a prepositional phrase. In each case, although the system of rules differs, the nature of the constraint is similar. That is, the rules determine which elements can legally combine to form structural units. Papers in this issue by Tanenhaus and Lucas and by Frazier discuss some of these different types of structural relations and their implications for the processes involved in spoken word recognition.

This type of structural context can be contrasted with contexts in which the relationship between elements does not result in some higher-level representation. One example of non-structural context is the associative relationship illustrated by the words *doctor* and *nurse*. These words are semantically related but they are not structurally related in the sense of forming a higher-level representation.

The distinction between structural and non-structural context effects is critical for certain autonomous models of lexical processing (e.g., Forster, 1981). To the extent that non-structural context effects can be explained in terms of relations holding between items within a *single* level of the processing system, they do not violate the autonomy hypothesis. In contrast, structural context effects involve *multiple* levels of the system and a top-down flow of information, and therefore are not consistent with strong versions of the autonomy view.

3.3. Non-structural context: Intra-lexical effects

The recognition of one word can have an impact upon the processing of another word which bears some relationship to the first. The relationship can be phonological, morphological, syntactic or semantic—but it does *not* involve the construction of a higher-level representation.

The intra-lexical context effects which are found for associatively related words provide a good example of non-structural context effect (Meyer & Schvaneveldt, 1971; Seidenberg et al., 1982). For example, the semantic properties of a word like *teacher* are assumed to prime or facilitate the processing of a semantically associated word such as *school*. Such priming effects have generally been interpreted within the framework of models like the Collins and Loftus (1975) spreading activation model of semantic networks. When the first word is presented, activation spreads to neighboring words or, rather, concepts with which it is linked. This activation translates into faster recognition of the related word.

Theorists like Forster and Fodor would argue that this type of "hardwired connection" between words is the only type of semantic context which can have an effect on any of the phases of word recognition. At issue then is the extent to which this type of context actually generalises to the recognition of words in utterance contexts, or whether it only operates when words appear in non-structured lists. Thus, the exact contribution of this type of context on lexical processing remains to be determined.

3.4. Structural context effects

3.4.1. Lexical context effects

Lexical effects refer to the influence that lexical representations are assumed to have upon acoustic-phonetic processing. Early evidence for such effects was furnished by Ganong (1980) who presented subjects with ambiguous phonemes (situated along a VOT continuum, e.g., k ↔ g continuum) in two different contexts. In one context, the first phoneme reading of the stimuli (e.g., /k/ in the context _iss) led to a word, whereas the other reading (/g/ in same context) produced a nonword (giss). Subjects gave more word phoneme responses than nonword responses, leading Ganong to argue that the lexical context within which phonemes are embedded influences the processing of these phonemes.

Two important questions must be answered in trying to characterise this effect: *how* and *when* are the two information sources brought into contact (cf. Segui & Frauenfelder, 1986)? With respect to the former question, we can distinguish two ways in which the higher level might affect the lower. Either it could contribute information to the lower level directly, or it could simply help in the evaluation of an autonomously constructed output.

In the Trace model (Elman & McClelland, 1984), lexical effects operate by modifying the analysis of the sensory input. This interactive activation model contains several levels of representation, each consisting of a set of interconnected nodes representing distinctive features, phonemes and words. These discrete, yet interactive, processing levels continuously exchange information. Incoming sensory input provides bottom-up excitation of the distinctive feature nodes which in turn activate phoneme nodes. As the phoneme nodes become excited, they can alter the level of activation of word nodes. Critically, the word nodes provide lexical feedback to lower level phoneme nodes, thereby increasing the level of activation of the phoneme node. As a consequence, phoneme processing and recognition depends not only on the bottom-up activation from the feature level but also on top-down influences from the lexical level. This view can be contrasted with an autonomous view in which the lexical level serves only to indicate the presence of a mismatch between its analysis and that of the lower level, and to suggest that a *revision* (or re-evaluation) of the analysis of the sensory input may be required.

In selecting further between these two alternative frameworks, we need to establish the temporal properties of lexical effects. First, we must determine the temporal locus of lexical effects with respect to: (a) the *phonetic categorisation* of the sensory input, and (b) *word recognition* (before or after a single lexical entry has been selected). Thus, it must be determined whether the lexicon influences phonetic decisions *before* or *after* phonetic categorisation is completed. Clearly, if a decision on the identity of a phonetic segment has already been made, lexical context can no longer have any effect upon the decision process itself, but only upon the process of evaluating the output of this process.

A complete account of the locus (or loci) of lexical effects must make reference not only to the temporal properties of acoustic-phonetic processing but also to those of lexical processing. In principle, the lexical level could exert its influence upon phonetic decisions either *after* word recognition (when a single lexical candidate has been selected) or *before* (when several lexical candidates are still active). Clearly, the contribution of the lexical level to phonetic processing is more valuable in the latter case. Indeed, only if the lexical level has its effect *during* the selection process of word recognition, can it really contribute to the process of recognising a word. If lexical context operates after word recognition, then its contribution is limited to serving other purposes—for example, identifying word boundaries.

Research into lexical effects has not yet provided adequate answers to the questions raised above. Nonetheless, despite the complexities of evaluating the contribution of lexical context to phonetic processing, the appropriate methodological tools appear to be within our grasp for addressing these questions empirically.

3.4.2. Syntactic context effects

A listener processing the syntactic structure of an utterance has information available that constrains the syntactic properties of upcoming words and their constituent structure. *How* and *when* does this syntactic information influence lexical processing? Theory (Frazier, 1987, this issue) and data (Tyler & Wessels, 1983; Tanenhaus et al., 1979) converge in attributing to syntactic context only a very limited capacity to intervene in the phases of lexical processing leading to word recognition.

This is not surprising since syntactic structure can impose only weak constraints on the form-class and morphological structure of lexical items. Although there is considerable cross-linguistic variation, in languages like English, it is only rarely possible to predict with absolute certainty the form-class of any given lexical item because each syntactic constituent has optional members. So, in most instances, a listener hearing the beginning of a noun phrase (e.g., having heard a determiner) cannot be certain of the form-class of the following item since it could be an adjective, adverb or noun.

However, there are rare cases where syntactic constraints do fix the form-class of a lexical entry. But, even in these cases their contribution still remains limited given the large number of words in each of the open class categories (e.g. adjectives, nouns, verbs). Moreover, in many instances the form-class of words is determined by their suffixes, or more precisely, their heads (Williams, 1981). Frazier (1987, this issue) argues that the fact that the rightmost derivational suffix generally determines a word's form-class limits the predictive power of the syntactic constraints. Listeners cannot confirm predictions based on preceding syntactic context until they correctly identify the element carrying the form-class information which often comes well after the stem has been heard. Some experimental evidence suggests that syntactic context can have an effect upon the recognition of suffixes. In a study examining the effects of different types of context on the recognition of polymorphemic words, Tyler and Marslen-Wilson (1986) found that syntactic constraints facilitate recognition of suffixes while semantic constraints facilitate recognition of stems. What this means is that the appropriate syntactic constraints can speed up the recognition of polymorphemic words, by facilitating recognition of suffixes. In these circumstances, a polymorphemic word may be recognised at the point at which its stem is recognised since the syntactic context can predict the suffixes.

3.4.3. Meaning-based context effects

We use the terms "semantic context" and "interpretative context" in order to highlight the distinction between a variety of types of contextual information which we believe ought to be distinguished, but rarely are. The term "semantic context" usually refers to any type of context which is meaningful—whether that meaning is based on such wildly differing types of meaning relations as semantic associations or pragmatic inference. In the interests of perspicuity, we propose that the term "semantic context" be reserved for the representation of an utterance which derives from the combination of those aspects of meaning which one wants to claim are based upon word meanings, together with the syntactic structure of the utterance. "Semantic context", then, would explicitly *not* involve pragmatic inferences, the discourse context and knowledge of the world. "Interpretative context" could then be used to refer to the more highly elaborated representation which incorporates these seemingly less tangible aspects of meaning (Marslen-Wilson & Tyler, 1980).

Semantic context. Unlike the effect of syntactic context, there is considerable evidence that semantic context affects lexical processing. Words which are semantically appropriate for the context are responded to faster and the identification of those which are contextually inappropriate are slowed down (Marslen-Wilson, 1984; Marslen-Wilson & Tyler, 1980; Morton & Long, 1976 (but see Foss & Gernsbacher, 1983); Tyler & Wessels, 1983). These facilitatory and inhibitory effects have been demonstrated using a variety of "on-line" tasks, for example, phoneme monitoring, lexical decision, mispronunciation detection, shadowing, naming, word monitoring and gating. There are two important issues in this research. First, there is the issue of the nature of semantic context effects. Second, there is the issue of where, in the various phases of lexical processing, these effects are to be located. Whenever we observe semantic context effects, we must determine whether they result from some type of higher-order structural representation of the words in an utterance or whether they can simply be explained in terms of intra-lexical associations. Because most researchers are neither explicit nor systematic about the type of meaning context they manipulate experimentally, the interpretation of their data is frequently difficult.

The difference between these two types of semantic context is critical for some autonomous models like that of Forster. While such models can accommodate intra-lexical semantic effects (because these can be located within a single level of the language system), they do not allow semantic context effects which are attributable to higher-level *structural* representations.

Undoubtedly, context effects of this latter type do exist, but what remains unresolved is exactly when they have their influence on lexical processing. If semantic context effects can be located in the early phases of lexical processing—as, for example, Grosjean (1980), Marslen-Wilson and Tyler (1980) and Morton (1969) would claim—they would be problematic for those autonomy models which confine structural information to a post-access role in lexical processing (Forster, 1979). Autonomy assumptions are not violated by such semantic context effects in Fodor's model because they can be located within the language module. It is only interpretative contextual effects which cannot be located within the language module and which involve the central processor (e.g., effects involving knowledge of the world) which violate autonomy assumptions.

Interpretative context. "Interpretative context" effects are certainly controversial for all versions of the autonomy thesis because such effects are outside the domain of an autonomous language processing system. If interpretative context can be shown to affect any of the phases of lexical processing prior to the point at which a single word is selected, then the autonomy assumption is violated. Such context effects are only consistent with the autonomy assumption if they operate after a single candidate has been selected.

Given the theoretical importance of this issue, we need to be able to demonstrate convincingly whether or not interpretative context does facilitate the processes involved in recognising spoken words. Much of the burden here is methodological. In addition to teasing apart the different types of meaning contexts distinguished above, we need to be able to specify what aspects of the word recognition process are reflected by different tasks. And we need to use tasks which are able, in principle, to tap the very *early* processes involved in word recognition. Otherwise there will be continuous dispute as to whether a particular task taps the appropriate phase of the process to allow us to make definitive claims about where context effects are to be located. As mentioned above, an attempt in this direction has been made recently with respect to lexical decision and naming tasks. Whether or not the current analyses of these tasks turn out to be correct, the enterprise is surely necessary.

4. Conclusions

We have focused in this short introduction on two major problems in the study of lexical processing—the phases involved in recognising a spoken word and the nature of different types of contextual influences on these phases. What we have attempted to do is to decompose the process of recognising a word into phases which have both theoretical and empirical consequences. There are necessarily pitfalls associated with such an enterprise. Although we found it necessary to break down lexical processing into its component phases, this does not mean that we assume that each phase is necessarily discrete. The extent to which the phases we differentiate are independent is still very much of an empirical issue. A related problem is terminological in nature; the terms we use to label these phases (like selection and lexical access) are ambiguous since they can refer either to a process or to the product of this process. A similar decompositional approach has been taken in our discussion of context effects. Here, we have attempted to distinguish qualitatively different types of context (lexical, intra-lexical, syntactic, semantic, and interpretative). This is necessary to make explicit the relationship between a particular type of top-down information and the phase(s) of processing at which it has its impact. Despite the problems inherent in the analytic approach taken here in this introduction, we consider it to be the most appropriate way to study lexical processing.

References

Aronoff, M. (1976). Word formation in generative grammar. *Linguistic Inquiry Monograph 1*. Cambridge, Mass.: MIT Press.

Becker, C.A., & Killion, T.H. (1977). Interaction of visual and cognitive effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 3*, 389–401.

Blank, M., & Foss D.J. (1978). Semantic facilitation and lexical access during sentence processing. *Memory and Cognition*, 6, 644–652.

Bradley, D.C., & Forster, K.I. (1987). A reader's view of listening. *Cognition, 25*, 103–134, this issue.

Bresnan, J. (1978). A realistic transformational grammar. In J. Bresnan, M. Halle & E. Miller (Eds.), *Linguistic theory and psychological reality*. Cambridge, Mass.: MIT Press.

Bybee, J. (1985). *Morphology. A study of the relation between meaning and form*. Amsterdam: Benjamins.

Cairns, H.S., & Hsu, J.R. (1980). Effects of prior context upon lexical access during sentence comprehension: A replication and reinterpretation. *Journal of Psycholinguistic Research, 9*, 319–326.

Church, K.W. (1987). Phonological parsing and lexical retrieval. *Cognition, 25*, 53–69, this issue.

Cole, R., & Jakimik, J. (1980). A model of speech perception. In R.A. Cole (Ed.), *Perception and production of fluent speech*. Hillsdale, N.J.: Erlbaum.

Collins, A., & Loftus, E. (1975). A spreading activation theory of semantic processing. *Psychological Review, 82*, 407–428.

Elman, J.L., & McClelland, J.L. (1984). Speech perception as a cognitive process: The interactive activation model. In N. Lass (Ed.), *Speech and Language, Vol. 10*. New York: Academic Press.

Fodor, J.A. (1983). *The modularity of mind: An essay on faculty psychology*. Cambridge, Mass.: MIT Press.

Ford, M., Bresnan, J., & Kaplan, R. (1982). A competence-based theory of syntactic closure. In J. Bresnan (Ed.), *The mental representation of grammatical relations*. Cambridge, Mass.: MIT Press.

Forster, K.I. (1976). Accessing the mental lexicon. In R.J. Wales & E. Walker (Eds.), *New approaches to language mechanisms*. Amsterdam: North-Holland.

Forster, K.I. (1979). Levels of processing and the structure of the language processor. In W.E. Cooper & E. Walker (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett*. Hillsdale, N.J.: Erlbaum.

Forster, K.I. (1981). Priming and the effects of sentence and lexical contexts on naming time: evidence for autonomous lexical processing. *Quarterly Journal of Experimental Psychology, 33A*, 465–495.

Foss, D.J. (1982). A discourse on semantic priming. *Cognitive Psychology, 14*, 590–607.

Foss, D.J., & Gernsbacher, M.A. (1983). Cracking the dual code: Toward a unitary model of phoneme identification. *Journal of Verbal Learning and Verbal Behavior, 22*, 609–632.

Fraunfelder, U.H. (1985). Cross-linguistic approaches to lexical segmentation. *Linguistics, 23*, 669–687.

Frazier, L. (1987). Structure in auditory word recognition. *Cognition, 25*, 157–187, this issue.

Ganong, W.F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance, 6*, 110–125.

Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception and Psychophysics, 28*, 267–283.

Grosjean, F., & Gee, J.P. (1987). Prosodic structure and spoken word recognition. *Cognition, 25*, 135–155, this issue.

Jakimik, J., Cole, R.A., & Rudnicky, A.I. (1985). Sound and spelling in spoken word recognition. *Journal of Memory and Language, 24*, 165–178.

Johnson-Laird, P.N. (1987). The mental representation of the meaning of words. *Cognition, 25*, 189–211, this issue.

Klatt, D.H. (1980). Speech perception: A model of acoustic-phonetic analysis and lexical access. In R.A. Cole (Ed.), *Perception and production of fluent speech*. Hillsdale, N.J.: Erlbaum.

Marcus, S.M., & Frauenfelder, U.H. (1985). Word recognition – uniqueness or deviation? A theoretical note. *Language and Cognitive Processes, 1–2*, 163–169.

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

Marslen-Wilson, W.D. (1984). Function and process in spoken word-recognition. In H. Bouma and D.G. Bouwhuis (Eds.), *Attention and Performance X: Control of language processes*. Hillsdale, N.J.: Erlbaum.

Marslen-Wilson, W.D. (1987). Functional parallelism in spoken word-recognition. *Cognition, 25*, 71–102, this issue.

Marslen-Wilson, W.D., & Tyler, L.K. (1980). The temporal structure of spoken language understanding. *Cognition, 8*, 1–71.

Marslen-Wilson, W.D., & Welsh, A. (1978). Processing interactions during word-recognition in continuous speech. *Cognitive Psychology, 10*, 29–63.

McClelland, J.L., & Rumelhart, D.E. (Eds.) (1986). *Parallel distributed processing: Explorations in the microstructure of cognition*. Cambridge, Mass.: Bradford Books.

Mehler, J. (1981). The role of syllables in speech processing: Infant and adult data. *Philosophical Transactions of the Royal Society, B 295*, 333–352.

Meyer, D., & Schvaneveldt, R. (1971). Facilitation in recognising pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90*, 227–234.

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review, 76*, 165–178.

Morton, J., & Long, J. (1976). Effect of word transitional probability on phoneme identification. *Journal of Verbal Learning and Verbal Behavior, 15*, 43–52.

Nooteboom, S. (1979). More attention for words in speech communication research. In B. Lindblom & S. Ohman (Eds.), *Frontiers of speech communication research*. London: Academic Press.

Norris, D.G. (1982). Autonomous processes in comprehension: A reply to Marslen-Wilson and Tyler. *Cognition, 11*, 97–101.

Norris, D.G. (1986). Word recognition: Context effects without priming. *Cognition, 22*, 93–136.

Pisoni, D.B., & Luce, P.A. (1987). Acoustic-phonetic representations in word recognition. *Cognition, 25*, 21–52, this issue.

Salasoo, A., & Pisoni, D.B. (1985). Interaction of knowledge sources in spoken word identification. *Journal of Memory and Language, 24*, 210–231.

Segui, J., & Frauenfelder, U.H. (1986). The effects of lexical constraints upon speech perception. In F. Klix & H. Hagendorf (Eds.) *Human memory and cognitive capabilities: Symposium in memoriam Hermann Ebbinghaus*, Amsterdam: North-Holland.

Seidenberg, M.S., Waters, G., Sanders, M., & Langer, P.S. (1984). Pre- and postlexical loci of contextual effects on word recognition. *Memory and Cognition, 12 (4)*, 315–328.

Seidenberg, M.S., Tanenhaus, M.K., Leiman, J.M., & Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology, 14*, 489–537.

Selkirk, E. (1984). *Phonology and syntax: The relation between sound and structure.* Cambridge, Mass.: MIT Press.

Shipman, D.W., & Zue, V.W. (1982). Properties of large lexicons: Implications for advanced isolated word recognition systems. *Proceedings of the 1982 IEEE International Conference on Acoustics, Speech and Signal Processing*, Paris, France, April.

Swinney, D. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior, 14*, 645–660.

Tanenhaus, M., Carlson, G., & Seidenberg, M. (1984). Do listeners compute linguistic representation? In D. Dowty, L. Karttunen & A. Zwicky (Eds.) *Natural language parsing.* Cambridge: Cambridge University Press.

Tanenhaus, M., Leiman, J., & Seidenberg, M. (1979). Evidence for multiple stages in the processing of ambiguous words in syntactic contexts. *Journal of Verbal Learning and Verbal Behavior, 18*, 427–441.

Tanenhaus, M.K., & Lucas, M.M. (1987). Context effects in lexical processing. *Cognition, 25*, 213–234, this issue.

Tyler, L.K. (1984). The structure of the initial cohort: evidence from gating. *Perception and Psychophysics, 26*, 417–427.

Tyler, L.K., & Marslen-Wilson, W.D. (1982). Conjectures and refutations: A reply to Norris. *Cognition, 11*, 103–107.

Tyler, L.K., & Marslen-Wilson, W.D. (1986). The effects of context on the recognition of multimorphemic words. *Journal of Memory and Language, 25*, 741–752.

Tyler, L.K., & Wessels, J. (1983). Quantifying contextual contributions to word-recognition processes. *Perception and Psychophysics, 34*, 409–420.

Tyler, L.K., & Wessels, J. (1985). Is gating an on-line task? Evidence from naming latency data. *Perception and Psychophysics, 38* (3), 217–222.

Williams, E. (1981). On the notions of "lexically related" and "head of a word". *Linguistic Inquiry, 12*, 245–274.

Résumé

Cette introduction a pour but d'élaborer le cadre général auquel appartiennent les articles rassemblés ci-dessous. Elle s'adresse à deux problèmes importants dans l'étude du traitement lexical: l'identification des différentes étapes dans la reconnaissance de mots et la caractérisation des divers types d'influences contextuelles sur ces étapes. Nous essayons de décomposer les processus de reconnaissance de mots en plusieurs étapes ayant des conséquences aux niveaux théorique et empirique. Nous adoptons également une approche analytique en traitant des influences dues au contexte, en distinguant plusieurs types de contexte (lexical, intra-lexical, syntaxique, sémantique et interprétatif). Une telle démarche est nécessaire si nous voulons rendre explicite le rapport entre tel ou tel type d'information contextuelle et l'étape où s'exerce son influence.