



Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Cognition 92 (2004) 101–144

COGNITION

www.elsevier.com/locate/COGNIT

The spatial and temporal signatures of word production components

P. Indefrey*, W.J.M. Levelt

Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, The Netherlands

Received 25 July 2001; revised 20 March 2002; accepted 26 June 2002

Abstract

This paper presents the results of a comprehensive meta-analysis of the relevant imaging literature on word production (82 experiments). In addition to the spatial overlap of activated regions, we also analyzed the available data on the time course of activations. The analysis specified regions and time windows of activation for the core processes of word production: lexical selection, phonological code retrieval, syllabification, and phonetic/articulatory preparation. A comparison of the word production results with studies on auditory word/non-word perception and reading showed that the time course of activations in word production is, on the whole, compatible with the temporal constraints that perception processes impose on the production processes they affect in picture/word interference paradigms.

© 2004 Elsevier B.V. All rights reserved.

Keywords: Language; Word production; Word perception; Neuroimaging

1. The functional organization of word production

Producing spoken words, whether in isolation or in the context of a larger utterance, involves an extensive neural network. In a recent meta-analysis of 58 neuroimaging studies of word production, Indefrey and Levelt (2000) found that the main components of this network, defined as cerebral regions showing statistically distinguishable activation in word production tasks, are largely left-lateralized. In addition, different regions of activation appeared to be involved with different functional components of the word production process. For instance, the conceptually driven selection of a lexical item, as in picture naming, typically goes with activation in the mid part of the left middle temporal

* Corresponding author.

E-mail addresses: indefrey@mpi.nl (P. Indefrey), pim@mpi.nl (W.J.M. Levelt).

gyrus. Similarly, the functional component of accessing a word's phonological code is linked to activation in Wernicke's area, whereas phonological encoding proper (syllabification and metrical encoding) shows corresponding activation in left inferior frontal regions.

The enterprise of relating the functional components of word production, such as lexical selection, phonological code retrieval, and syllabification, to regions in a cerebral network requires a detailed, explicit theory of the process of spoken word production. The meta-analysis presented here is based on the theory presented in Levelt, Roelofs, and Meyer (1999), henceforth to be called LRM. The theory explicates the successive computational stages of spoken word production, the representations involved in these computations, and their time course. The results of the meta-analysis, however, do not hinge on this particular choice of theory, since differences between the sequential LRM model and other models of word production (Butterworth, 1989; Dell, 1986; Dell, Schwartz, Martin, Saffran, & Gagnon, 1997; Garrett, 1980; Stemberger, 1985) do not concern the assumed processing levels but the exact nature of the information flow between them. The method and design of the neuroimaging experiments analyzed here were not suited to identify these rather subtle differences between current models.¹

LRM is largely based on evidence from chronometric experiments. The nature of these chronometric experiments allows us to address two further issues. The first one concerns the time course of activation within the lexical cerebral network. Voice onset latencies in word production studies not only provide evidence about the total duration of lexical preparation (for instance from picture onset to the initiation of the naming response – typically some 600 ms) but also about further temporal details of this preparation. In a picture/word interference experiment, for instance, a distracter word is auditorily or visually presented while the subject is preparing the picture's name. The distracter can be presented at different stimulus onset asynchronies (SOAs), before, simultaneous with, or after picture onset. The effects of distracters on the naming latencies vary with SOA. Typically, a temporal order effect is obtained. Semantic distracters (for instance *goat* when the target name is *sheep*) have their maximum effect on response latencies at shorter SOAs than phonological distracters (for instance *sheet* when the target word is *sheep*). Such data provide additional information about the time windows for semantically driven selection and phonological preparation of the target word. It is a challenge to relate such functional time course information to the time course of cerebral activation in word production imaging studies.

The second issue concerns the relations between the cerebral word production and word perception networks. Although theories of word production and word perception have been developing in rather independent research traditions, there cannot be any reasonable doubt that the two functional systems are intimately linked. This is already apparent from the chronometric paradigms used in word production studies. The core paradigm of picture/word interference demonstrates the effectiveness of linguistic input in affecting

¹ Considering that no two tasks used in the neuroimaging experiments analyzed here differed with respect to the involvement of the conceptual and lemma processing levels, the results of the meta-analysis are also compatible with the assumption that these two processing levels should not be distinguished (Caramazza, 1997; Starreveld & La Heij, 1996).

the time course of word production. Semantic or phonological stimuli systematically inhibit or facilitate aspects of lexical access in word production. Although we still lack detailed, experimentally supported theories of the functional integration of word perception and production, some minimal assumptions can be made about this integration which are compatible with the mentioned experimental evidence from word production studies. Here the challenge is to relate these ‘linking assumptions’ to the activations observed in purely perceptual imaging studies that use the same types of semantic and phonological stimuli as input.

The following two sections will elaborate these two issues, after which we will turn to the relevant imaging evidence.

1.1. Components and time course of word production

The LRM theory conceives the production of spoken content words as a staged process, beginning with selecting some target lexical concept for expression, and ending with the initiation of articulation. Each stage generates its own characteristic output representation and has a distinct time course. The staging architecture is depicted in the two rightmost columns of Fig. 1. The stages correspond to strata in an activation spreading network, called *WEAVER* (Roelofs, 1992, 1997): a conceptual stratum in which nodes represent lexical concepts, i.e. concepts for which there are words in the target language, a syntactic or lemma stratum in which nodes represent the syntax of words, and a form stratum where nodes represent various form properties of words. The staging of lexical access in spoken word production, from conceptual preparation, via lemma retrieval, to form encoding, corresponds to activation spreading from stratum to stratum in the lexical network. Lexical access proper is followed by the motor execution stage of articulation. Let us consider these stages and their time course in turn.

1.1.1. Conceptual preparation

Producing a content word normally starts by activating some lexical concept and by selecting it for expression. For instance, when you are asked to name a picture, you must recognize the depicted object and select an appropriate concept. A long tradition of speech error and picture naming research (see Levelt, 1999 for a review) has demonstrated that there is normally multiple activation of lexical concepts in response to visual input. The picture of a sheep not only activates the concept *SHEEP*,² but probably also concepts such as *ANIMAL* or *GOAT*. It depends on the communicative situation or the experimental task which concept is going to be selected for expression. In a categorization task, for instance, the subject will select the superordinate concept (*ANIMAL*), and in a normal naming task it is usually the basic level concept (*SHEEP*). This selection strategy is called ‘perspective taking’ (Clark, 1997; Levelt, 1989, 1996).

The speed of selecting the target concept is, of course, not a fixed quantity. The zooming in on a target for expression is a rhetorical decision, which may take any amount of rumination on the part of the speaker. Still, latencies in a standard picture naming

² Following the notational tradition, we will denote lexical concepts in capital letters, lemmas in italics, and phonological codes by way of IPA symbols.

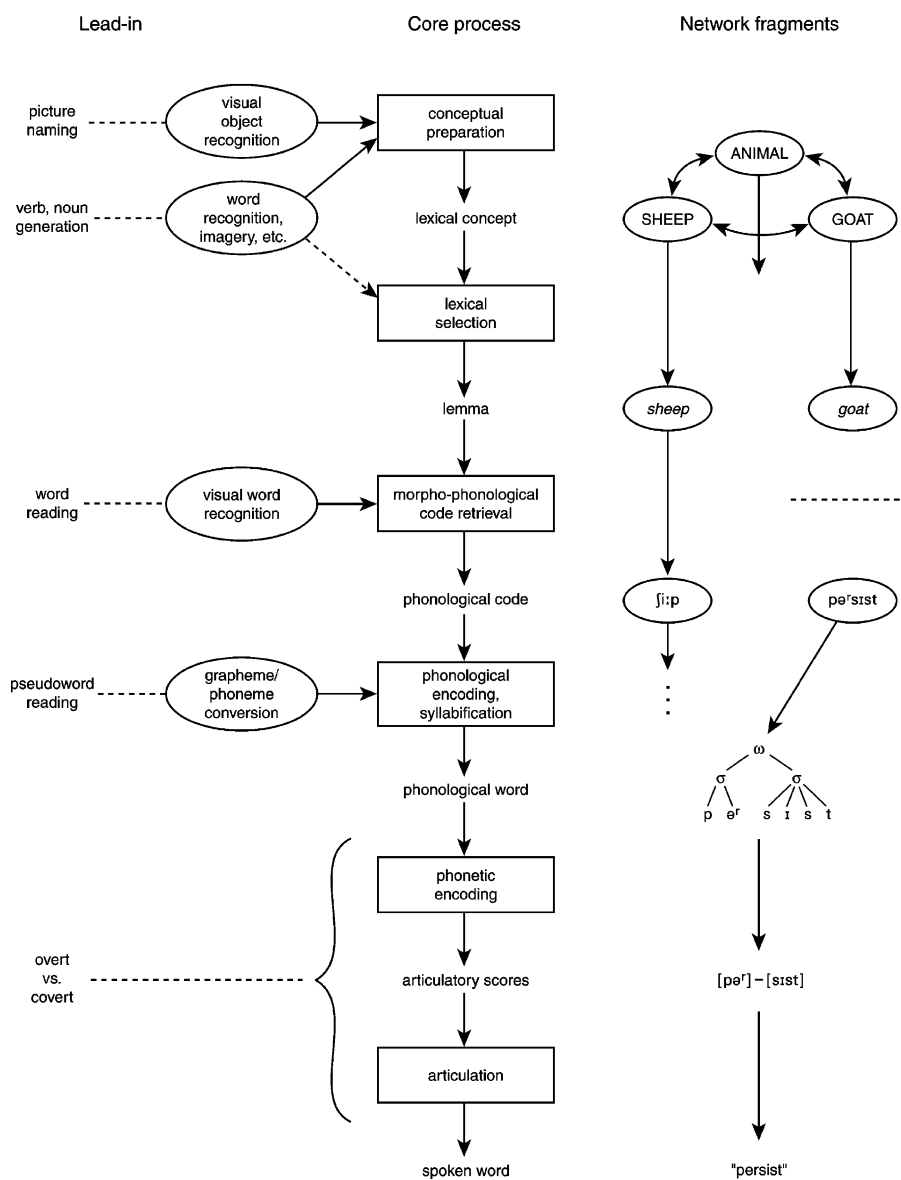


Fig. 1. Processing network and componential task analysis. Left column: experimental tasks and their 'lead-in' processes. Middle column: core processes of word production and their characteristic output. Right column: example fragments of the WEAVER spreading activation network and its output.

paradigm show moderate variability, allowing us to estimate the time window for selecting the target concept. Thorpe, Fize, and Marlot (1996), in a go/no-go event-related potential (ERP) study, asked subjects to categorize a presented picture as animate or non-animate. The evoked electrical scalp potentials obtained for the two kinds of stimuli started to

diverge around 150 ms after picture onset. The notion ANIMATE was apparently activated within 150 ms. This corresponded to a push button response time of 445 ms. Preparing and executing the response apparently took some 300 ms. In order to check access to the lexical concept, Jescheniak and Levelt (1994) performed a picture recognition task similar to Thorpe et al.'s. The subject first saw a word on the screen denoting the target (for instance *sheep*), then saw a picture. If the picture corresponded to the target (i.e. depicted a sheep), the subject pushed the 'yes' button, otherwise the 'no' button. This task requires accessing the lexical concept (i.e. SHEEP). The average 'yes' response occurred 439 ms post picture onset, a latency closely corresponding to that obtained by Thorpe et al. If again some 300 ms are spent on response preparation and execution, accessing the lexical concept must have occurred within a 150 ms time frame. Schmitt, Münte, and Kutas (2000) report slightly longer latencies. Like Thorpe et al. (1996) they used an animacy decision task, but here a dual paradigm was used. In the relevant condition the animacy of the depicted object determined whether a push button response should be made; it governed a go/no-go decision. The response itself (i.e. in the go condition) concerned the target word's initial phonological segment: vowel versus consonant. Recorded ERPs (at midline sites) showed an N200 effect, an increased negativity for no-go trials as compared to go trials. The average onset of the N200 effect (defined as 10% of the peak effect) occurred at 206 ms post picture onset. In other words, at that moment the subject had accumulated enough evidence about the animacy of the depicted object to withhold responding. Of course, the 'withhold' response itself also needs time to develop. Hence, the 206 ms reflect an upper limit on accessing the animacy concept. Summarizing the evidence so far, accessing the lexical concept from a visual object appears to take place within a time frame of 150–200 ms post picture onset. Our median estimate is 175 ms.

1.1.2. Lemma retrieval (lexical selection)

The next stage involves accessing the target word's syntax. In normal utterance production the most urgent operation after conceptual preparation is the incremental construction of a syntactic frame, i.e. grammatical encoding. Word order, constituent (part-of-speech) formation, and inflection all depend on the syntactic properties of the lexical items that are accessed. Lemma nodes in the syntactic stratum of the lexical network represent these syntactic properties (such as word category, gender of nouns, syntactic argument structure of verbs). How is a lemma node selected? As mentioned, perspective taking typically involves multiple activation of lexical concepts, only one of which gets selected for expression. In the model, each node at the conceptual stratum is linked to its unique lemma node at the syntactic stratum. If multiple concepts are activated, multiple lemma nodes will get co-activated. The latency of selecting the target lemma decreases with its degree of activation and increases with the degree of activation of co-activated alternative lemmas. Roelofs' (1992) mathematical competition model for lemma selection has found quantitative support in a host of experimental studies (see Levelt et al., 1999 for a review). In this model, the parameter settings accounting optimally for the semantic interference effects reported by Glaser and Döngelhoff (1984) for different SOAs provided an estimate of the time window for lemma selection of 100–150 ms. This corresponds well to the 115 ms estimate derived from modelling the dual task data in

Levelt et al. (1991). Here the subjects' standard task was one of picture naming. But on one-third of the trials the subject was presented with an auditory target for lexical decision. If the target was a word, it could be semantically or phonologically related to the picture's name, or unrelated. The target could be presented at various SOAs. The lexical decision latencies revealed the expected time course from early semantic interference to later phonological interference. The semantic interference reflects the activation of the picture name's target lemma after object recognition. In a recent ERP study Schmitt, Schiltz, Zaake, Kutas, and Münte (2001) directly addressed the lemma access time window. In one condition subjects were instructed to respond to the picture only in case the depicted object was lighter than 500 g. In the 'go' case, they had to choose a response hand on the basis of the (German) gender of the target picture's name. The latter is a syntactic, lemma property of the picture's name. In another condition these two conditional responses were reversed: the gender determined go versus no-go and the conceptual information 'weight' determined the response hand. In both conditions an N200 arose, i.e. an early negativity in the no-go situation (as compared to the go situation). The N200 peaked at 477 ms when the go/no-go decision depended on weight; it peaked at 550 ms when the decision depended on gender. In the LRM framework the 73 ms difference reflects an upper boundary of the time window for retrieving a noun's lemma, given its lexical concept, since suprathreshold activation of the gender node may be subsequent to lemma selection. Given that no lower boundary is available, a conservative estimate for the range of lemma selection durations (in response to picture stimuli) is between 0 and 150 ms. Our median estimate is 75 ms. Given the above estimate for accessing the lexical concept, 150–200 ms post picture onset, the operation of lemma selection should begin between 150 and 200 ms post picture onset and be over at some moment between 150 and 350 ms post picture onset.

1.1.3. Form encoding

The range of operations involved in form encoding begins with accessing the target word's phonological code and ends while the word is being articulated. Clearly, the encoding of a word's onset should be complete at the moment articulation is initiated. An estimate for the total time frame of form encoding can be obtained by subtraction. In LRM accessing the target word's phonological code starts upon selection of the lemma. According to the above estimates, lemma selection is complete between 150 and 350 ms post picture onset. When we subtract this from the average picture naming latency, we have an upper bound on the duration of encoding the word-initial articulatory program. Most picture/word interference experiments do not provide usable estimates of picture naming latencies, because of the presence of a distracter stimulus. Undistracted picture naming latencies are reported by Jescheniak and Levelt (1994) (680 ms), by Levelt, Praamstra, Meyer, Helenius, and Salmelin (1998) (591 ms), and by Damian, Vigliocco, and Levelt (2001) (567 ms).³ The average estimate is 600 ms. As in most picture/word interference experiments, these studies involved repeated exposure of the same pictures during the experiment. First time naming of a picture is often substantially slower.

³ Data are taken from averages over low and high frequency conditions in the former two cases (Dutch language), and the 'heterogeneous' (i.e. low semantic interference) condition in the latter case (German language).

The range of average latencies, from 567 to 680 ms, is partly due to word frequency. LRM locates the word frequency effect (you are relatively slow for low-frequency words, usually words that are acquired later in life) in the operation of accessing the word's phonological code. Measured word frequency effects range from 0 (Levelt et al., 1998) to about 60 ms (Jescheniak & Levelt, 1994). If lemma selection is complete between 150 and 350 ms in this type of experimental setting, subtraction leaves us with a duration range for form encoding from 217 to 530 ms. The wide range of this estimate is obviously due to the cumulative effect of uncertainties in a subtraction procedure.

Form encoding, however, is itself a staged process and there is some independent evidence for the durations of these substages of form encoding. According to LRM, the first operation upon lemma selection is morphological encoding (morpho-phonological code retrieval). The speaker accesses the phonological codes for all of the target word's morphemes. For instance, when the lemma *goat* has the syntactic diacritical feature *pl*, indicating that the target is the plural form of the word, then two phonological codes will be retrieved, one for the stem and one for the plural inflection (/gout/ and /s/, respectively). For the irregular word 'sheep', only one code will be retrieved, /ʃi:p/. For details of morphological encoding and its chronometry, see Janssen, Roelofs, and Levelt (2002). So far, imaging studies of word production never systematically varied morphological complexity. In other words, in these studies the complexities of morphological encoding reduce to 'accessing the phonological code'. In LRM accessing a morpheme's code is essentially accessing its (possibly underspecified) phonological segments. Van Turenout, Hagoort, and Brown (1998) provided ERP evidence indicating a rather short latency for retrieving a word's phonological code. Their lateralized readiness potential (LRP) measurements indicate that a word's first phonological segment (the word-initial consonant) is available at about 40 ms after the word's syntactic gender information has been accessed. If we grant that lemma selection might precede the availability of gender information by up to 73 ms, as discussed in the previous section, the estimated range for the availability of the first phonological segment is between 40 and 113 ms after lemma selection. Our median estimate is around 80 ms.

The second operation in form encoding is phonological encoding proper. For spoken word production, this reduces to syllabification and metrical encoding. In LRM syllabification is an incremental process. The 'spelled-out' segments of the phonological code are incrementally clustered in syllabic patterns. For instance, to syllabify the verb 'persist', you first cluster the initial segments /p/ and /ə/ to create the first syllable /pə/; then you cluster the next four spelled-out segments to form the word's second syllable /sɪst/, with the syllabified phonological word /pə-sɪst/ as a result. Syllabification is not fixed in the lexicon, but produced 'on-line' in a context-dependent fashion. For instance, the bimorphemic progressive version of the same verb will be syllabified as /pə-sɪs-tɪŋ/, where the second syllable is /sɪs/, not /sɪst/. This segment-by-segment internal syllabification proceeds at a speed of about 25 ms per segment, as appears from LRP experiments (Van Turenout, Hagoort, & Brown, 1997) and phoneme self-monitoring experiments, in which reaction times increased at about this rate as a function of the word internal position of phonemes and interacted with syllable structure (Wheeldon & Levelt, 1995). On this estimate, an average five-segments-word will need some 125 ms to syllabify.

The third operation is phonetic encoding. As syllables are incrementally created, they are rapidly turned into motor action instructions. In LRM these instructions ('syllable

Table 1
Estimated time windows for successive operations in spoken word encoding

Operation	Duration (ms)
Conceptual preparation (from picture onset to selecting the target concept)	175
Lemma retrieval	75
Form encoding:	
Phonological code retrieval	80
Syllabification	125
Phonetic encoding (till initiation of articulation)	145
Total	600

scores') are stored for the few hundred high-frequency syllables that do most of the work in normal speech production. The repository of articulatory syllable scores is called the 'mental syllabary'. We have no independent chronometric measures of phonetic encoding. Estimates about its duration can only be obtained by subtraction (with the obvious proliferation of uncertainty). An additional problem here is that articulation of a plurisyllabic word can be initiated before completion of phonetic encoding. This was convincingly argued by [Bachoud-Lévi, Dupoux, Cohen, and Mehler \(1998\)](#). Minimally, the first syllable must have been phonetically encoded before articulation can be initiated. [Meyer, Roelofs, and Levelt \(2003\)](#) showed that the amount of phonetic (but not of phonological) encoding before the initiation of articulation is a strategic decision on the part of the speaker, not a 'hard-wired' constant.

[Table 1](#) presents our best estimates for phase durations in picture naming, where the average naming latency is put at 600 ms. The noticed variability of the estimates should caution against too rigid interpretation of these numbers.

1.2. Some relations between word production and word perception

The functional unity of the speaker/hearer suggests close integration of the language user's perceptual and production networks. However, in spite of much suggestive evidence in the literature, a focused research endeavour to study this integration experimentally never materialized. The perception and production of speech are still largely approached independently. The evidence for their mutual relations mainly comes from 'mixed' experimental paradigms within the one or the other research tradition and we will not go much beyond that in the present study. Still, some of the evidence raises important issues for the interpretation of neuroimaging data in word processing. The relevant relations are depicted and numbered in [Fig. 2](#). The success of the classical picture/word interference paradigm in word production research shows the effectiveness of linguistic input in affecting the word production mechanism. LRM accounts for these effects by a threefold connection between the networks (the numbers correspond to those in [Fig. 2](#)).

(1) Distracter words that are semantically related to the spoken word target (for instance 'goat' when the target is 'sheep') typically slow down the naming response. [Roelofs](#)

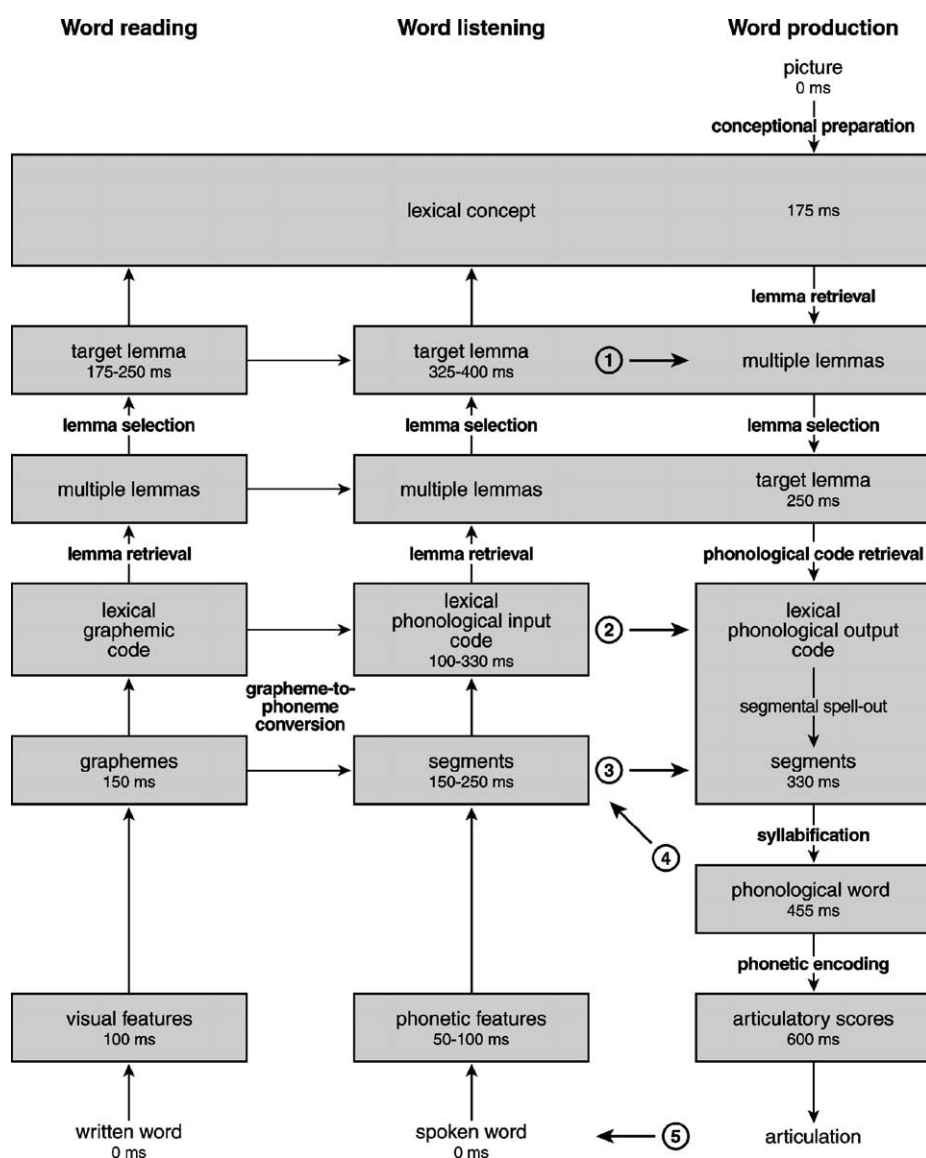


Fig. 2. Network of processing components involved in speech production and perception. Left column: assumed processing steps in word reading. Middle column: assumed processing steps in word listening. Right column: core processes of word production. Numbered arrows 1–5 indicate interactions of production and perception pathways (see text).

(1992) could handle all existing SOA curves by assuming that the distracter word, whether spoken or written, affects the state of activation of the corresponding lemma in the production network. LRM is blank on the precise perceptual mechanism mediating between the perceptual input and the resulting lemma activation. However, importantly,

LRM assumes that the word production and perception networks are shared from the lemma level ‘upwards’, i.e. for the lemma stratum and the conceptual stratum. A lemma activated by perceptual input spreads its activation to the corresponding lexical concept. Because production requires the reverse direction of activation spreading, from concept to lemma, the connections between these two strata should be bi-directional.

(2) Distracter words, whether spoken or written, that are phonologically related to the target word typically speed up the naming response (relative to neutral controls). LRM assumes that this manipulation directly affects the state of activation of the morpheme nodes, i.e. the phonological codes in the production network. The mechanism, however, is left unspecified in LRM (and it will be different for spoken and written word input). Still, it is the obvious channel mediating the fast reading of a printed word. The visual input activates the relevant phonological code. From there on the normal core production process (syllabification → phonetic encoding → articulation) is run. Notice that LRM rejects the notion that phonological codes are shared between perception and production. They may be closely linked, but they are not identical. Various arguments for this claim are discussed in [Levelt et al. \(1999\)](#). Hence, there is no direct ‘perceptual’ feedback from phonological production codes to lemmas (but see (4) below). Although the direct, fast reading route does not involve lemma activation, route (1) does allow for lemma activation (and hence conceptual activation) in reading. And that is how it should be: we sometimes do read for meaning.

(3) There is substantial experimental evidence for spoken word production priming by non-word spoken or visual distracters (reviewed in [Levelt et al., 1999](#)). LRM assumes that such non-morphemic perceptual input directly affects the state of activation of the corresponding segment nodes in the production network. Again, the mechanism is left unspecified.

(4,5) Self-monitoring is an essential property of the speaking architecture. [Levelt \(1983\)](#) explained part of the speaker’s self-monitoring ability by assuming the involvement of the speech perception system. Speakers can self-perceive their overtly or internally produced speech: the ‘external’ (5) and ‘internal’ (4) perceptual loop, respectively. [Wheeldon and Levelt \(1995\)](#) provided evidence that the internal loop takes the (string of) phonological words as input, i.e. the output of phonological encoding, not of phonetic encoding. Mechanisms (1) through (3) above suffice to account for these feedback loops ([Roelofs, in press](#)). Self-produced internal or overt speech are natural ‘distracters’. Overt self-produced words activate the corresponding lemmas via route (1), the corresponding phonological codes via route (2), and the corresponding phonological segments via route (3). It is a small step to assume that the internal monitoring loop functions by taking phonological (i.e. pre-phonetic) words as input to the perceptual network, further following routes (1) through (3).

The proposed connections between the two networks, together with the time course information discussed in the previous section, impose theoretical restrictions on the time course of the relevant perceptual processes. For a semantically related spoken word to affect lemma selection in picture naming, the perceptual network must have activated the distracter lemma within the time window of lemma retrieval. Written word distracters show their maximal effect at SOAs of about 0 ms. In other words, perceptual, printed-word lemma activation must peak within the 175–250 ms time window (post word

presentation). Spoken word distracters require negative SOAs to be maximally effective. Schriefers, Meyer, and Levelt (1990) obtained the strongest effect at $SOA = -150$ ms. This is at least partly due to the incremental character of the distracter stimulus. Adding 150 ms to the just estimated time window, lemma activation in spoken word perception should peak within a 325–400 ms time window post word onset.

According to LRM phonological facilitation from word distracters involves both routes (2) and (3). Such distracters facilitate access to the target word's phonological code and its segments in a time window between 250 and 330 ms post picture onset. Basic data for spoken word distracters were reported by Meyer and Schriefers (1991). The optimal SOAs were either 0 or 150 ms, dependent on whether initial or final parts of the distracter word facilitated the target phonological code (see Roelofs, 1997 for a detailed analysis). In other words, phonological code activation by a spoken word distracter should peak between 100 ms (250 minus 150 ms) and 330 ms post word onset. The range, however, may again be substantial. The effect of non-word distracters involves route (3) alone, but there is no evidence that the time window of this effect differs from the phonological word effects just discussed. Phonological priming can also be induced by means of written stimuli, whether words or non-words. Most studies do not provide SOA data (Ferrand, Segui, & Grainger, 1996; Lupker, 1982; Schiller, 1998). Essentially all effects in these studies were obtained at a single SOA. Most commonly the (masked) distracter appeared 40–60 ms before the picture. On LRM's assumption that such distracters also affect the spell-out of the phonological code, a printed word or non-word is still 'phonologically active' in the critical time range between 290 ms (250 plus 40 ms) and 390 ms (330 plus 60 ms) post stimulus presentation. It could, however, be active much earlier. This was in fact shown in two studies, which do provide SOA data for orthographic phonological priming. Starreveld and La Heij (1996) used unmasked visual prime words, which shared the target word's initial (C)V. They found substantial priming effects for an SOA range of -200 to $+100$ ms. Damian and Martin (1999) used the same SOA range in their unmasked visual priming experiment and obtained essentially the same results. Hence, an orthographic prime can be effective over a broad range of 150 ms (250 minus 100 ms) to 530 ms (330 plus 200 ms) post prime onset. However, the range becomes substantially smaller for masked visual primes. When Damian and Martin repeated the experiment using short (200 ms) masked visual primes, significant effects were only observed for the SOA range of 0 to $+200$ ms, and the effects were smaller. The range of SOAs corresponds to an effective time range for phonological facilitation by written distracters of 50–330 ms post prime onset. In other words, the size and temporal scope of orthographic priming depends strongly on the conspicuity of the prime.

Any imaging effect of (post-lexical) self-monitoring minimally requires the availability of the target word's phonological representation, i.e. the output of phonological encoding. Since phonological encoding is an incremental process, its output is also an incremental product. Self-monitoring can probably start as soon as the word-initial phonological syllable has been composed, i.e. some 25 ms after retrieval of the word's phonological code. Following Table 1, this is around 355 ms post picture onset. But then, self-monitoring can continue till after the word has been overtly spoken; there is no clear limit at the far end.

2. Cerebral regions involved in word production and perception – a meta-analysis

The meta-analysis by Indefrey and Levelt (2000) identified a set of cerebral areas related to the core processes of word production. In the present paper, the focus is on cerebral areas that are possibly involved in the interaction of word production and perception. This means that in addition to the set of areas involved in word production we seek to identify areas that are active both in word production and perception. To this end, the present meta-analysis covers a large number of studies using two word production tasks involving the complete stream of production processes (picture naming and word generation), two word production tasks that start with visual word perception and enter the production processing stream at later stages (word and pseudoword reading), and two auditory perception tasks (listening to words and pseudowords). We will further examine whether the timing information that is available for areas that are activated by production and perception processes is compatible with the time windows during which – according to chronometric data – interactions occur.

2.1. A component analysis of word production and perception tasks

From the perspective of word production, Indefrey and Levelt (2000) distinguished between core processes and so-called lead-in processes of word production tasks. Lead-in processes are task-specific cognitive processes, such as visual object recognition in picture naming, taking place before the core word production pathway is entered. These processes are not well understood for all tasks, but they always contribute essentially to the neuroimaging results. Some lead-in processes involved in the four word production tasks analyzed here are listed in the second column of Fig. 1. Note, however, that without serious behavioural research, one can only speculate at the processes involved in some task lead-ins.

The four word production tasks differ not only in their lead-in processes, but they also enter the cascade of “core” processes of word production at different levels (Fig. 1, third column). Consequently, they do not share the same core processing components.

2.1.1. Picture naming

In picture naming the task enters the componential hierarchy from the very top component, conceptual preparation. The lead-in process is visual object recognition, which provides an object percept as input to conceptual preparation. Although this lead-in process is quite well understood, still many variables are to be controlled, such as visual complexity, perspectival orientation of the object, colour versus black-and-white and of course object category. All core components of word production are involved in picture naming.

2.1.2. Verb generation

In this task, the subject is asked to generate one or more appropriate verbs to a given noun (APPLE → “eat”). Similar to picture naming, this task also involves all core components of word production. However, the lead-in process is ill-understood (cf. Indefrey, 1997). The subject sees or hears a noun, which triggers a visual or auditory word recognition process. If the noun is a concrete one, the subject will probably generate

a visual image, which under the perspective of the task activates one or more associated actions in long-term memory. These, then, guide the further conceptual preparation. When the noun is abstract, long-term memory may get accessed without visual imagery. But there are possible shortcuts, too. A perceived noun may directly activate a verbal concept or even occasionally a verb lemma by sheer association, like in *knife–cut*.

2.1.3. Noun generation

The typical task here is to present a semantic category, such as “jobs” or “tools” or “animals”, and the subject is asked to generate as many exemplars as possible. It is a so-called “word fluency” task. The lead-in process may involve something as complicated as an imaginary tour, such as mentally touring a zoo. It may also be a lower-level process, such as word association. And the subject’s strategy may differ rather drastically for different semantic categories. It is quite likely, though, that at least from lexical selection onwards all core processes of word generation are involved.

2.1.4. Word reading

The lead-in process is visual word recognition, which is complicated enough by itself (see below). The core process may either start at the level of syllabification, from a set of activated phonemes (the phonological route), or it may start from the activated lexical entry providing the phonological code. The relative contributions of sublexical orthographic-to-phonological conversion and lexical access from the graphemic code depend on context, word frequency, and orthographic regularity, which may differ from language to language, even from word to word. Since in most studies reviewed here direct lexical access from the visual word form was possible for at least a part of the stimuli, we assume activation of the lexical phonological code to be the first core processing component of word production.

2.1.5. Pseudoword reading

Here no syntactic word or a word’s phonological code are accessed, but there is syllabification. The lead-in process is visual orthographic analysis, some kind of bottom-up grapheme-to-phoneme mapping, which provides the ordered pattern of phonemes as input to syllabification.

Insofar as written words or pseudowords function as distracters in picture naming, we are interested in the kind of phonological or semantic information that becomes available through visual orthographic analysis. As depicted in Fig. 2, we assume processing steps at the level of visual features, graphemes, lexical graphemic codes, and lemmas. It is, however, beyond the scope of this article to provide a functional model of reading. In particular the distinction between single graphemes and lexical graphemic codes is highly simplified. Most probably sublexical phonological recoding operates on grapheme strings of variable sizes. Note that Fig. 2 is also not meant to represent a particular view on the degrees of parallel processing and interactivity in visual orthographic analysis.

2.1.6. Word and pseudoword listening

Similarly, we do not attempt to provide a functional model of spoken word processing (for a review of spoken word comprehension models see Cutler & Clifton,

1999). We assume that spoken words are processed at the level of phonetic features, segments, lexical phonological codes and lemmas. Spoken word distracters may affect word production at the lexical (phonological code or lemma) and the sublexical level.

2.2. *Meta-analysis procedures*

It is the purpose of this section to identify the neural substrates of the different processing components that have been laid out in the previous section. To this end, we will analyze the data reported in a large number of studies according to the following heuristic principle: if, for a given processing component, there are subserving brain regions, then these regions should be found active in all experimental tasks sharing the processing component, whatever other processing components these tasks may comprise. The region(s) should not be found active in experimental tasks that do not share the component.

This approach allows for the isolation of processing components between studies even if isolation within single studies is not possible. Nevertheless, four conditions must be met. First, the processing components must be independently defined, so that their absence or presence can be evaluated for every experiment by applying the same criteria (which may differ from the author's criteria). Second, the task and control conditions must be heterogeneous enough across different experiments to secure that a specific processing component is the only component that is shared. Thirdly, the task and control conditions must be heterogeneous enough across different experiments to ensure that for every processing component there is a different set of tasks that share the component. Finally, the database must be large enough to comprise enough experiments for a reliable identification of activations typically found for the different tasks. Indefrey and Levelt (2000) have shown that these requirements can be met for word production. This holds a fortiori for the present paper, since the database of word production experiments has been considerably extended.

The two word perception tasks, passive listening to words and pseudowords, have been applied in neurocognitive research frequently enough to allow for a reliable identification of the associated network of cerebral areas. By taking into account timing data we will attempt to identify areas within this network that may be related to interactions between spoken word distracters and word production. For a comprehensive discussion of the neural correlates of the processing components of auditory word comprehension, we refer the reader to the articles by Richard Wise and Sophie Scott and by Dana Boatman in this volume.

2.2.1. *Data set*

We analyzed the localization data from 82 word production experiments and 26 auditory perception experiments (Table 2). To identify the cerebral regions involved in the experimental tasks, the majority of studies used methods detecting hemodynamic changes, such as positron emission tomography (PET), functional magnetic resonance imaging (fMRI), or single photon emission computed tomography (SPECT). The remaining studies used subdural or direct cortical electrical stimulation, transcranial magnetic stimulation (TMS), lesion data, or magnetoencephalographic (MEG) recordings. In order to be able to

Table 2
Overview of experiments included in the data set

Study	Method	Production	Control condition
<i>Picture naming</i>			
Ojemann (1983)	Cortical, thalamic stimulation	Overt	
Ojemann, Ojemann, Lettich, and Berger (1989)	Cortical stimulation	Overt	
Schäffler, Lüders, Dinner, Lesser, and Chelune (1993)	Cortical stimulation	Overt	
Haglund, Berger, Shamseldin, Lettich, and Ojemann (1994)	Cortical stimulation	Overt	
Salmelin, Hari, Lounasmaa, and Sams (1994)	MEG	Overt	
Abdullaev and Melnichuk (1995)	Single cell recordings	Overt	Blank screen
Bookheimer, Zeffiro, Blaxton, Gaillard, and Theodore (1995)	PET	Covert	Nonsense drawing
Damasio, Grabowski, Tranel, Hichwa, and Damasio (1996)	PET	Overt	Faces “up, down”
Damasio et al. (1996)	Lesion data	Overt	
Kiyosawa et al. (1996)	PET	Overt	Rest
Malow et al. (1996)	Subdural electrical stimulation	Overt	
Martin, Wiggs, Ungerleider, and Haxby (1996)	PET	Covert	Nonsense objects
Price, Moore, Humphreys, Frackowiak, and Friston (1996)	PET	Covert	Object viewing “yes”
Martin, Wiggs, and Weisberg (1997)	PET	Covert	Visual noise
Levelt et al. (1998)	MEG	Overt	
Zelkowitz, Herbst, Nebes, Mintun, and Becker (1998)	PET	Overt	Nonsense objects, “hiya”
Chao and Martin (1999)	PET	Overt	Grey scale Mondrians
Moore and Price (1999)	PET	Overt	Meaningless objects, “okay”
Murtha, Chertkow, Beauregard, and Evans (1999)	PET	Overt	Plus sign
Etard et al. (2000)	PET	Overt	Rest
Van Turennout, Ellmore, and Martin (2000)	fMRI	Covert	Visual noise
Hamberger, Goodman, Perrine, and Tamny (2001)	Cortical stimulation + subdural grids	Overt	
Stewart, Meyer, Frith, and Rothwell (2001)	TMS	Overt	
<i>Word generation</i>			
Wise et al. (1991)	PET	Overt	Rest
McCarthy, Blamire, Rothman, Gruetter, and Shulman (1993)	fMRI	Overt	Rest

(continued on next page)

Table 2 (continued)

Study	Method	Production	Control condition
Eulitz et al. (1994)	PET	Covert	Rest
Crivello et al. (1995)	PET	Overt	Rest
Poline, Vandenberghe, Holmes, Friston, and Frackowiak (1996): 12 Exps.	PET	Overt	Rest
Warburton et al. (1996): Exps. 1B, 2B, 3A, 3B, 4B	PET	Overt	Rest
Paulesu et al. (1997)	fMRI	Overt	Rest
Warburton, Price, Swinburn, and Wise (1999)	PET	Covert	Rest
Etard et al. (2000)	PET	Overt	Rest
Papathanassiou et al. (2000)	PET	Covert	Rest
<i>Word reading</i>			
Petersen, Fox, Posner, Mintun, and Raichle (1989)	PET	Covert	Fixation
Petersen, Fox, Snyder, and Raichle (1990)	PET	Covert	Fixation
Howard et al. (1992)	PET	Overt	False fonts, “crime”
Price et al. (1994)	PET	Covert	False fonts
Price et al. (1994)	PET	Overt	False fonts, “absent/present”
Bookheimer et al. (1995)	PET	Covert	Nonsense drawings
Bookheimer et al. (1995)	PET	Overt	Nonsense drawings
Menard, Kosslyn, Thompson, Alpert, and Rauch (1996)	PET	Covert	xxXxx
Price, Moore, and Frackowiak (1996)	PET	Covert	Rest
Price, Moore, and Frackowiak (1996)	PET	Overt	Rest
Beauregard et al. (1997)	PET	Covert	Expecting words, fixation
Herbster, Mintun, Nebes, and Becker (1997)	PET	Overt	Letterstrings, “hiya”
Martin et al. (1997)	PET	Covert	Visual noise
Rumsey et al. (1997)	PET	Overt	Fixation
Brunswick, McCrory, Price, Frith, and Frith (1999)	PET	Overt	Rest
Fiez, Balota, Raichle, and Petersen (1999)	PET	Overt	Fixation
Hagoort et al. (1999)	PET	Covert	Fixation
Moore and Price (1999)	PET	Overt	False fonts, “okay”
Tarkiainen, Helenius, Hansen, Cornelissen, and Salmelin (1999)	MEG	Covert	
Cohen et al. (2000)	fMRI	Covert	Rest
Mechelli, Friston, and Price (2000)	PET, fMRI	Covert	Rest
Salmelin, Schnitzler, Schmitz, and Freund (2000)	MEG	Overt	
Tagamets, Novick, Chalmers, and Friedman (2000)	fMRI	Covert	Geometric shapes
Veltman, Friston, Sanders, and Price (2000)	PET	Covert	False fonts

(continued on next page)

Table 2 (continued)

Study	Method	Production	Control condition
<i>Pseudoword reading</i>			
Petersen et al. (1990)	PET	Covert	Fixation
Herbster et al. (1997)	PET	Overt	Letterstrings, “hiya”
Martin et al. (1997)	PET	Covert	Visual noise
Rumsey et al. (1997)	PET	Overt	Fixation
Brunswick et al. (1999)	PET	Overt	Rest
Fiez et al. (1999)	PET	Overt	Fixation
Hagoort et al. (1999)	PET	Covert	Fixation
Mechelli et al. (2000)	PET, fMRI	Covert	Rest
Tagamets et al. (2000)	fMRI	Covert	Geometric shapes
Indefrey, Hagoort, Herzog, Seitz, and Brown (2001)	PET	Overt	False fonts
<i>Word listening</i>			
		Type of stimuli	
Petersen et al. (1989)	PET	Words	Fixation
Price et al. (1992)	PET	Words	Rest
Mazoyer et al. (1993)	PET	Words	Rest
Binder et al. (1994)	fMRI	Words	Scanner noise
Fiez et al. (1995)	PET	Words	Fixation
Binder, Frost, Hammeke, Rao, and Cox (1996)	fMRI	Words	Scanner noise
Fiez, Raichle, Balota, Tallal, and Petersen (1996)	PET	Words	Fixation
Malow et al. (1996)	Subdural electrical stimulation	Words	
Mellet et al. (1996)	PET	Words	Rest
Price, Wise et al. (1996): Exps. 4 and 5	PET	Words	Rest
Warburton et al. (1996)	PET	Words	Rest
Dhankhar et al. (1997)	fMRI	Words	Scanner noise
Maddock and Buonocore (1997)	fMRI	Words	Scanner noise
Rees et al. (1997)	PET, fMRI	Words	Rest, scanner noise
Cardebat, Démonet, Puel, Agniel, Viallard and Celsis (1998)	SPECT	Words	Rest
Mummary, Ashburner, Scott, and Wise (1999)	PET	Words	Noise
Wise, Greene, Büchel, and Scott (1999)	PET	Words	Anticipation
Wong, Miyamoto, Pisoni, Sehgal, and Hutchins (1999)	PET	Words	Rest
Kuperberg et al. (2000)	fMRI	Words	Rest
<i>Pseudoword listening</i>			
Wise et al. (1991)	PET	Pseudowords	Rest
Binder et al. (1994)	fMRI	Pseudowords	Scanner noise
Fiez et al. (1995)	PET	Meaningless CV syllables	Fixation
Fiez, Raichle et al. (1996)	PET	Pseudowords	Fixation
Shtyrov, Kujala, Ilmoniemi, and Näätänen (1999)	MEG	Meaningless CV syllables	
Binder et al. (2000)	PET	Pseudowords	Rest

identify the core processes of word production and perception, we applied the following selection criteria.

We included production experiments of four tasks, i.e. picture naming, word generation, word reading, and pseudoword reading. The set of word generation experiments comprised both verb generation (finding one or more semantically appropriate verbs to a given noun) and noun generation (finding nouns of a prespecified semantic category) experiments.

Reading tasks have both a production and a perception component. In order to keep the data set homogeneous with respect to the processing of the written input, only experiments involving the reading of Roman script were included in the database.

We did not include production experiments reporting enhanced cerebral activations during word production tasks relative to control tasks which themselves comprised most or all of the word production process, for example, reading aloud (Abdullaev & Posner, 1997; Buckner, Raichle, & Petersen, 1995; Fiez, Raichle et al., 1996; Petersen et al., 1989; Raichle et al., 1994; Snyder, Abdullaev, Posner, & Raichle, 1995) or object naming (Martin, Haxby, Lalonde, Wiggs, & Ungerleider, 1995). Our approach did also not allow for the inclusion of experiments or task comparisons focussing on the relative strengths of components of the word production process, for example comparisons of reading regularly versus irregularly spelled words (Herbster et al., 1997). Activations of these two tasks relative to baseline, however, were included. It was assumed throughout that the reported activation foci reflected true increases during the tasks rather than decreases during the baseline conditions.

For auditory perception we included studies involving passive listening to words or passive listening to pseudowords. Meaningless CV-syllables presented in two studies (Fiez et al., 1995; Shtyrov et al., 1999) were considered as monosyllabic pseudowords. Tasks involving additional decision processes, such as phoneme monitoring in pseudowords or semantic monitoring in words (cf. Démonet et al., 1992), were excluded.

2.2.2. Anatomical coding

The reported activation foci were coded in a descriptive reference system of 110 regions. In this system, the cerebral lobes were divided into two or three rostro-caudal or medio-lateral segments of roughly equal size. The segment labels were defined in terms of Talairach co-ordinates as given in Table 3. The regions within this gross division were defined in terms of gyri and subcortical structures following Talairach and Tournoux (1988). Cingulate, insular, and cerebellar activations were further differentiated descriptively (see Table 4). Activation foci located near the border of two adjacent regions were coded in both regions.

2.2.3. Reliability estimate

The studies included in this meta-analysis were not given any weights reflecting reliability differences due to design or size. This means that a certain degree of overlap of activations between studies was considered reliable, but should not be interpreted as statistically significant. Nonetheless, the notion of ‘reliability’ was not totally arbitrary, but based on the following quasi-statistical estimate: the average number of activated regions per experiment divided by the number of regions equals the probability for any

Table 3
Definition of descriptive anatomical labels: subdivisions of cortical lobes

Frontal		Temporal		Parietal		Occipital	
Anterior	$y > 34$	Anterior	$y > -7$	Sensory (approx.)	$y > -23$	Medial	$ x \leq 25$
Posterior	$34 \geq y \geq 0$	Mid	$-7 \geq y \geq -38$	Anterior	$-23 \geq y \geq -48$	Lateral	$ x > 25$
Motor (approx.)	$y < 0$	Posterior	$y < -38$	Posterior	$y < -48$		

particular region to be reported in an experiment if reports were randomly distributed over regions. Assuming this probability, the chance level for a region to be reported as activated in a number of experiments is given by a binomial distribution. We rejected the possibility that the agreement of reports about a certain region was coincidental if the chance level was less than 10%.

Assuming, for example, an average number of 11 activated regions per study, this reliability criterion corresponded to a minimum agreement of two studies for regions covered by two to five studies, a minimum agreement of three studies for regions covered by six to 11 studies, and so forth (four out of 12–18; five out of 19–25). In this way, the reliability threshold controlled for the fact that due to the heterogeneity of techniques and analysis procedures not all studies covered the whole brain. Note that for regions covered by many studies a relatively smaller number of positive reports was required to be above chance (comparable to five times ‘6’ with ten dice throws happening less easily by chance than one ‘6’ with two dice throws). The procedure also controlled for the fact that for some tasks typically more regions are found activated than for others, so that the chances of coincidental agreements of findings between studies increase.

We validated the procedure by applying it to the 12 data sets of a multi-centre study on verb generation reported by Poline et al. (1996). Among 22 regions judged reliable on the basis of the above criterion there were no false positives compared to the pooled statistical analysis performed by the authors. Our procedure rejected 17 regions that were statistically significant in the pooled analysis but found in less than four of the 12 single-centre analyses. Note that the 12 experiments analyzed by Poline et al. (1996) did not only use the same task, but were also standardized with respect to the experimental procedures. It is probable that regions showing significant activation in a small number of single experiments had similar trends in others that contributed to the pooled analysis results. We

Table 4
Definition of descriptive anatomical labels: subdivisions of cingulum, insula, and cerebellum

Cingulum		Insula		Cerebellum	
Anterior	$y > 12$	Anterior	$y > 0$	Medial	$ x \leq 20$
Mid	$12 \geq y \geq -24$	Posterior	$y \leq 0$	Lateral	$ x > 20$
Posterior	$y < -24$				

therefore considered it appropriate to use a more conservative criterion for the more heterogeneous data sets of the present meta-analysis.

In consequence, the reliability criterion we applied does not mean that atypical findings of activations in any single study are necessarily coincidental. In many cases, the number of experiments *not* reporting activations was not sufficient to consider a region as inactive at the chosen error probability level. Rare observations do, therefore, not exclude the possibility that a region is active. They may, for example, reflect smaller activations that are only detectable with refined techniques or better scanning devices. A second, related point is that the nature of the data does not allow for an interpretation in terms of relative strengths of activations of certain areas. It is known that parameters such as item duration and frequency strongly influence the resulting pattern of activations (Price, Moore, & Frackowiak, 1996; Price et al., 1994). It is thus possible that areas are more frequently found active in some tasks, because their ‘typical’ item durations and frequencies are higher or lower than in other tasks. It seems wise not to overinterpret the data, given that there is a considerable variability of these parameters across the studies of our database; also, the interactions of these parameters with other experimental factors are largely unknown.

2.2.4. Terminology

Combining data from different techniques made it necessary to find a common term for cerebral localizations observed in relation to certain tasks. Since the majority of experiments used PET or fMRI we use the terms ‘activations’ or ‘activated areas’, and extend that use to MEG sources and to sites where certain functions are interfered with by cortical stimulation or lesions. We are aware that for the latter case one can at best infer that such locations are ‘active’ in normal functioning.

2.3. Overall results

In a first analysis step, data from all 82 word production experiments were collapsed. The average number of subjects was 11.8 and a mean of 12.4 activated regions were found per experiment. Due to the large number of experiments, 40 reliably not activated regions could be identified that were found no more than three times in 60 or more experiments (see Fig. 3 and Appendix A). These regions were clustered in the superior and medial parietal lobe, the right anterior and medial frontal lobe, and the anterior inferior temporal lobes bilaterally. In addition, the posterior cingulate, the hypothalamus, and the hippocampus (all bilateral) were rarely reported for word production tasks.

2.4. Neural correlates of task-specific lead-in processes

Both picture naming and word generation recruit all word production processing components. Therefore, activation areas that were found reliably for only one of the two tasks can be assumed to be related to task-specific processes rather than to the core processing components of word production. Given that word generation was performed silently in all experiments but one of the data set, whereas the majority of picture naming experiments involved overt articulation, the two tasks differed not only with respect to

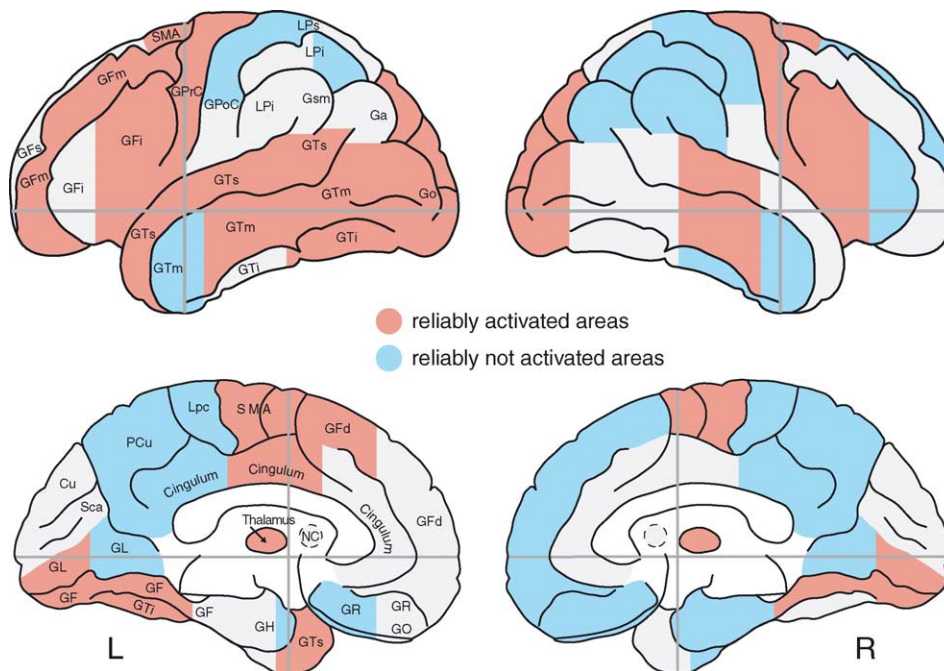


Fig. 3. Schematic lateral (top) and medial (bottom) views of the brain indicating reliably activated and not activated cerebral regions based on 82 word production experiments (R, right hemisphere; L, left hemisphere). Note that insula, cerebellum, and most subcortical structures are not depicted in this brain schema. The intersection of the horizontal and vertical grey lines indicates the origin (anterior commissure) of the coordinate system of Talairach and Tournoux (1988). The abbreviations of gyri and subcortical structures follow Talairach and Tournoux (1988) except for SMA, supplementary motor area. Cu, Cuneus; Ga, angular gyrus; GF, fusiform gyrus; GFd, medial frontal gyrus; GFs, GFm, GFi, superior, middle, and inferior frontal gyrus; GH, parahippocampal gyrus; GL, lingual gyrus; GO, orbital gyri; Go, occipital gyri; GPoC, postcentral gyrus; GPrC, precentral gyrus; GR, gyrus rectus; Gsm, supramarginal gyrus; GTs, GTm, GTi, superior, middle, and inferior temporal gyrus; LPs, LPi, superior and inferior parietal lobule; NC, caudate nucleus; PCu, precuneus; Sca, calcarine sulcus.

their lead-in processes but also with respect to the processes of phonetic encoding and articulation. Hence, to study the lead-in processes of picture naming, activations should be taken into account that were specific for picture naming when compared to word generation (Appendix B) and at the same time were not specific for overt responses in general (Appendix C). Twelve such regions were reliably found (Fig. 4, green regions): six left and right occipital areas, the left mid temporal fusiform gyrus, the right posterior temporal fusiform gyrus, left mid and posterior sections of the inferior temporal gyrus, the right posterior inferior frontal gyrus, and the mid cingulate. All of these regions have been reported for tasks involving object viewing, the principle lead-in process of picture naming (Bookheimer et al., 1995; Kosslyn, Alpert, & Thompson, 1995; Kosslyn et al., 1994; Martin et al., 1996; Moore & Price, 1999). Visual processing functions are furthermore suggested by the observation that most of these regions were also reliably found for the two reading tasks involving visual stimulus presentation.

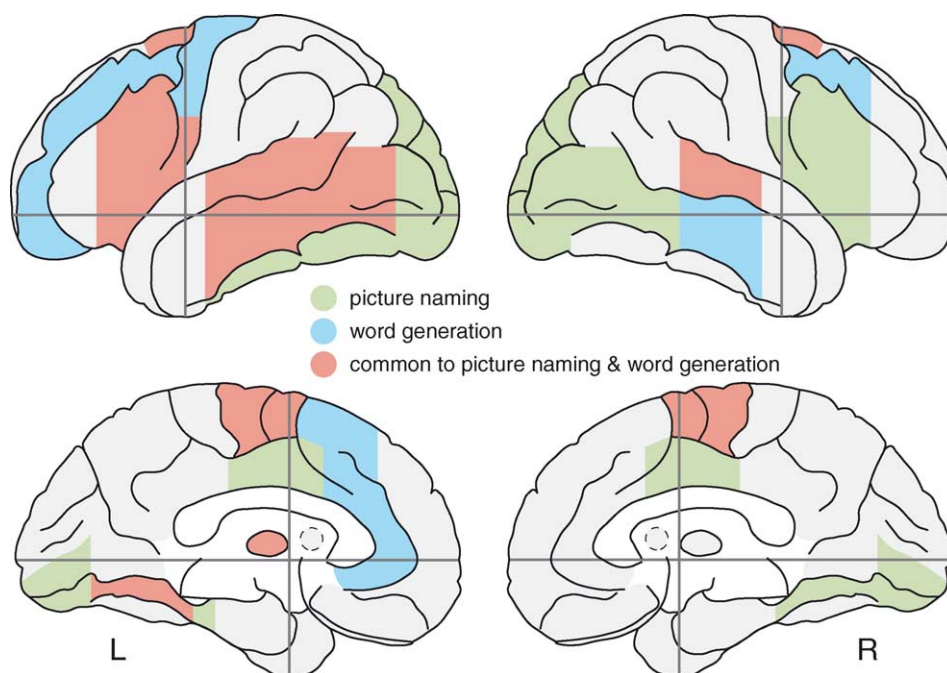


Fig. 4. Reliable regions for picture naming and word generation. Regions shared by both tasks are assumed to be involved in the core process of language production. Additional regions related to word generation were the right anterior insula and the left lentiform nucleus. Additional regions common to word generation and picture naming were the right medial and lateral cerebellum, the left medial cerebellum, and the left anterior insula.

Eight areas were reliably found in word generation tasks but not in picture naming (blue regions in Fig. 4): the left anterior cingulate, the right anterior insula, the left lentiform nucleus, the left dorsal precentral gyrus, the left anterior and posterior (bilaterally) middle frontal gyri, and the left posterior medial frontal gyrus. Extensive prefrontal and anterior cingulate activations seem to be specifically related to lead-in processes of word generation. Fiez, Petersen, and Raichle (1996) found activations of these regions, when comparing word generation to repetition of the stimulus word. In the same comparison, Raichle et al. (1994) and Fiez, Petersen, and Raichle (1996) observed blood flow *decreases* in the insulae. The authors suggest that the kind of response selection subserved by the insulae is different from the generation-specific response selection. In this case, however, the reliable insular activations in the generation tasks would have to be attributed to the auditory processing of the stimulus words. This interpretation is not supported by the data on passive word listening (see Section 2.6).

Both the left posterior medial frontal and dorsal precentral regions are immediately adjacent to the left supplementary motor area (SMA). Although left SMA activations were also reliably found in the other production tasks, they were more frequently reported for word generation. It is conceivable that large SMA activations in word generation extend into neighbouring regions. On the other hand, a word generation-specific activation of

the dorsal medial frontal gyrus, which might increase chances for the detection of adjacent SMA activations, cannot be excluded.

2.5. Neural correlates of the core processes of word production

Picture naming and word generation are the two tasks that include all component processes of the word production stream. The set of regions that were reliably found in both tasks (red regions in Fig. 4) can be considered as being related to the core process of word production. This word production network consisted of four right-hemispheric regions (mid superior temporal gyrus, medial and lateral cerebellum, and SMA) and 11 left-hemispheric regions (posterior inferior frontal gyrus, ventral precentral gyrus, SMA, mid and posterior superior and middle temporal gyri, posterior temporal fusiform gyrus, anterior insula, thalamus, and medial cerebellum) (Appendix B). Compared to the results of the meta-analysis of Indefrey and Levelt (2000), who found only six common areas for picture naming and word generation, this is a considerable increase reflecting the enhanced sensitivity of the larger database. The six previously found areas (left posterior inferior frontal gyrus, left mid and posterior superior and middle temporal gyri, and left thalamus) are confirmed in the present analysis. The increase is due to regions that according to the previous meta-analysis were either related to overt responses but now passed the threshold also for covert word generation (left ventral precentral gyrus and bilateral medial cerebellum), or were task-specific (bilateral SMA, right mid superior temporal gyrus, left anterior insula, left fusiform gyrus, right lateral cerebellum).

By taking into account word and pseudoword reading, which enter the word production processing stream at later stages, we now attempt to identify the subprocesses, to which each of the 15 regions may be particularly sensitive.

2.5.1. Conceptual preparation and lexical selection

In word production, the selection of a lemma is a conceptually driven process, whereas in reading it is part of the perception process. It is, therefore, assumed that word generation and picture naming but not word reading should share regions subserving the conceptually driven lexical access. As in Indefrey and Levelt (2000), only one region, the mid section of the left middle temporal gyrus, showed this pattern. Recent timing evidence (see Section 3) suggests a role for this region in lexical selection rather than conceptual processing as such. Considering the evidence about widespread and possibly category-specific areas being involved in prelexical conceptual processing (Beauregard et al., 1997; Damasio et al., 1996; Martin, 2000; Martin et al., 1995, 1996; Vandenberghe, Price, Wise, Josephs, & Frackowiak, 1996), it seems plausible that tasks like picture naming and word generation, which probably activate quite different concepts, should only converge and enter a common pathway from the point of lexical selection onwards.

2.5.2. Phonological code retrieval

Lexical word form retrieval takes place in picture naming, word generation, and word reading, but not in pseudoword reading. This pattern was found in the reported activations of the right SMA, the left anterior insula, and the left posterior superior and middle temporal gyri (Wernicke's area). While the left anterior insula and the right SMA did not

show the same pattern in Indefrey and Levelt (2000),⁴ the present meta-analysis confirms the earlier findings with respect to Wernicke's area. This was also found for the subset of pseudoword reading studies that had not been included in the previous meta-analysis (Brunswick et al., 1999; Fiez et al., 1999; Martin et al., 1995; Mechelli et al., 2000; Tagamets et al., 2000). Wernicke's area was also reliably found in passive word listening experiments (see Section 2.6) suggesting that it may serve as a common store of lexical word form representations for word production and perception.

2.5.3. Syllabification

All production tasks involve the cascade of word production processes from syllabification onwards. Syllabification is conceived of as operating on an abstract segmental representation, whereas in the subsequent stages of phonetic encoding and articulation motor representations are built up and executed. In the case of covert language production the latter processes obviously stop at a certain point, but this point depends on the exact nature of the covert task. Typical task instructions range from "thinking" over "covertly articulating" to "mouthing" of responses. For a neural correlate of syllabification, therefore, reliable activation in all production tasks is a necessary but not a sufficient condition, since motor areas, too, were found in all tasks. It seems, nonetheless, reasonable to assume that regions involved in the planning of articulatory movements are more frequently found activated for overtly pronounced as compared to covert responses. Regions involved in syllabification, by contrast, should not show any sensitivity to the overt–covert distinction. In a collapsed data set across all tasks comparing overt experiments involving silent control conditions with covert experiments (Appendix C), only the left posterior inferior frontal gyrus (Broca's area) met this requirement. Six other regions that were reliably reported in all production tasks (left ventral precentral gyrus, bilateral mid superior temporal gyri, left posterior temporal fusiform gyrus,⁵ left thalamus, and right medial cerebellum) were found at least twice as frequently in overt production experiments as compared to covert production. Across tasks, even the left SMA, which was most frequently found in covert word generation (see above), showed a higher probability to be reported in overt experiments. This result confirms the result of Indefrey and Levelt (2000) with respect to a role of Broca's area in syllabification. A role of the left mid superior temporal gyrus in this process, suggested by the results of the previous meta-analysis, has become less probable considering the sensitivity of this region to overt responses.

2.5.4. Phonetic encoding and articulation

Due to the following three considerations, the regions involved in phonetic encoding and articulation were most difficult to identify. Firstly, as discussed in the previous section, the data suggest that phonetic/articulatory planning took place not only in the case of overt responses but also in the case of covert responses. Secondly, experiments involving overt

⁴ Considering the sensitivity of the right SMA to the distinction between overt and covert responses, this region seems to be involved in articulatory planning rather than lexical phonological code retrieval (see below). For the left anterior insula, too, a role in articulatory planning has been suggested (Dronkers, 1996). This, however, is not supported by the overt–covert comparison showing no major difference for the left anterior insula.

⁵ The relatively smaller number of reports of left posterior temporal fusiform activation in covert tasks is mainly due to the fact that this area was less frequently reported in silent word generation.

and covert response were not equally distributed across tasks. More specifically, experiments involving overt responses were more likely to also involve visual stimulus presentation (pictures or word-like stimuli). Thirdly, overt responses are perceived as spoken input and used for self-monitoring, so that some regions that were preferentially activated in overt response experiments may be involved in auditory perception rather than articulation. Taking these difficulties into account, we assumed areas to be related to phonetic encoding and articulation that were (a) reported at least twice as frequently for overt as compared to covert responses (see ‘proportion overt’ and ‘proportion covert’ in Appendix C), (b) relatively more frequently reported for overt responses across tasks than in any of the tasks involving visual processing, and (c) not reliably found in word perception.

Seventeen such regions were found, of which 12 (bilateral ventral motor and sensory regions, right dorsal motor region, right SMA, left and medial right cerebellum, bilateral thalami, right midbrain) are known to be part of central nervous motor systems. It is plausible that these areas should contribute to the planning and execution of articulatory movements. The sensitivity of five other regions (right posterior inferior frontal gyrus, left orbital gyrus, bilateral posterior lingual gyri, and right posterior medial temporal fusiform gyrus) to the overt–covert distinction was unexpected and may be due to a modulatory effect of overt naming in visual processing tasks (Friston et al., 1996).

2.5.5. *Self-monitoring*

Self-monitoring involves an internal loop, taking as input the phonological word, i.e. the output of syllabification (see arrow 4 in Fig. 2), as well as an external loop, taking as input the acoustic speech signal of the own voice (see arrow 5 in Fig. 2). There is evidence that hearing one’s own voice while speaking induces the same temporal lobe activations as listening to someone else’s voice (McGuire, Silbersweig, & Frith, 1996; Price, Wise et al., 1996). We can therefore assume regions to be involved in the external loop of self-monitoring if they were found reliably in the word listening task and were more strongly activated in experiments involving overt responses. This was the case for the bilateral superior temporal gyri with the exception of the right anterior section. An involvement of the bilateral superior temporal gyri in the external loop of self-monitoring is supported by data from McGuire, Silbersweig, and Frith (1996) and Hirano et al. (1997), who were able to induce additional bilateral superior temporal activations by distorting the subjects’ feedback of their own voice or presenting the subjects with alien feedback while they spoke.

The tasks included in this meta-analysis do not enable the identification of cerebral regions subserving the internal loop of self-monitoring. The most economical assumption is that this loop also enters the pathway that is used for speech perception. In this case, the areas involved in internal self-monitoring would be either identical to those identified for the external loop, or a subset of these areas, depending on when the speech perception pathway is entered. McGuire, Silbersweig, Murray et al. (1996) provided some evidence that internal monitoring makes use of the left posterior superior temporal lobe. This area showed stronger activation (together with motor and premotor areas) when subjects imagined hearing another person’s voice than when they spoke silently to themselves. It does not seem implausible that the observed blood flow increase was due to an attentional modulation of internal self-monitoring, although other explanations are possible as well. If one accepts the premiss that auditory hallucinations are based on some

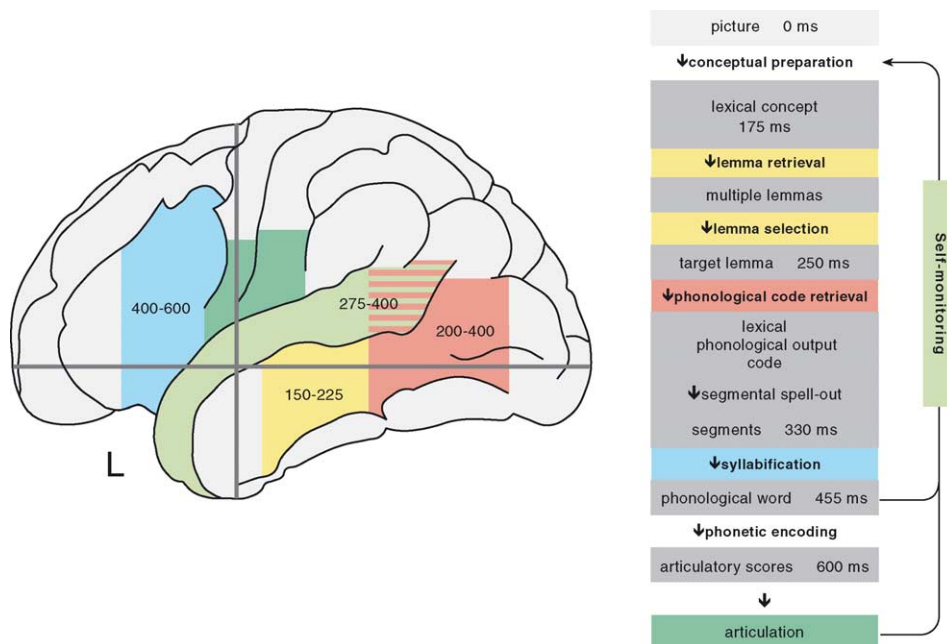


Fig. 5. Left column: schematic representation of meta-analysis results for word production. Identical colours indicate relations between regions and functional processing components (right column). The numbers indicate the time windows (in milliseconds) during which the regions are activated in picture naming (see text in Section 3). Further regions involved in phonetic encoding and articulation are the right sensorimotor cortex, the right SMA, the left and medial right cerebellum, the left and right thalamus, and the right midbrain. A further region involved in self-monitoring is the right mid superior temporal gyrus. Right column: time course of picture naming as estimated from chronometric data.

alteration of normal internal self-monitoring (David & Busatto, 1999; McGuire et al., 1995), evidence from this rather different line of research may be taken into account, too. On the one hand, mid and posterior superior temporal areas have been found activated during auditory hallucinations (Dierks et al., 1999; Lennox, Park, Medley, Morris, & Jones, 2000; Shergill, Brammer, Williams, Murray, & McGuire, 2000). On the other hand, the response of the superior temporal cortex to external auditory input seems to be reduced during hallucinations (David et al., 1996).

In sum, the results of the meta-analysis suggest that within the network of regions subserving the core process of language production there is functional specialization. Fig. 5 summarizes the tentative assignments of functional roles discussed in the preceding sections.

2.6. Shared neural correlates of word production and perception processes

Passive listening to words and pseudowords showed largely overlapping activation patterns. With the exception of the posterior part of the right middle temporal gyrus, all

regions of the superior and middle temporal gyri were reliably found activated (see Appendix B). In word listening experiments, furthermore, activation of the left posterior inferior frontal gyrus was reliably found.

Comparing the set of regions activated during auditory word perception to the set of regions that were reliably found in word production tasks, there were eight common regions: the bilateral mid and posterior superior temporal gyri, the left anterior superior temporal gyrus, the left mid and posterior middle temporal gyrus, and the left posterior inferior frontal gyrus. Among these, two groups should be distinguished. As discussed above, the bilateral superior temporal gyri were preferentially reported in word production experiments with overt responses, suggesting some function in self-monitoring. By contrast, the three remaining regions (mid and posterior middle temporal gyrus, left posterior inferior frontal gyrus) were insensitive to the overt–covert distinction and have been identified as being involved in three subsequent stages of the word production pathway: lexical selection, lexical phonological code retrieval, and post-lexical syllabification. Thus, at least as far as the spatial overlap between the neural correlates of word production and perception processes is concerned, this result is not incompatible with the psycholinguistic evidence suggesting that points of contact between the word production and perception pathways are located at all three stages (see Section 1.2 and Fig. 2).

3. The time course of cerebral activations during word production

So far, we have seen that neural structures subserving core processing components of word production are also activated in word perception. This spatial overlap may play a role in a neural explanation for interactions between the two pathways, for example by minimizing relay times for the exchange of information. Temporal overlap, however, is much more crucial. Psycholinguistic effects like the interference of semantically related distracter words with the selection of the target lemma in word production require a precise timing overlap between the pathways. The perceived information must be available at the right time and in the right format to affect the production of spoken words.

3.1. Time windows of word production components

We will first examine whether the tentative assignment of functional roles in Section 2.5 is supported by the time courses observed in MEG experiments of picture naming. MEG measurements of the magnetic fields induced by the electrical activity of neurons do not only provide information on the location of magnetic sources but also have a temporal resolution in the order of milliseconds. Therefore, they provide a link between the spatial activation patterns identified in the meta-analysis and the time windows of the component processes of word production that were estimated in Section 1 on the basis of chronometric data, such as reaction time experiments and ERP studies. The chronometric data serve as independent predictors. They will be compared to the MEG activation time courses of the regions that, according to the meta-analysis, are related to the different processing

components. If the regions indeed subserve the assumed processes, their activation time courses must be compatible with the chronometric data. In the following section, we will then examine whether the available MEG data on word listening and reading allow for the identification of neural correlates of interactions between word perception and word production.

3.1.1. Conceptually driven lexical selection

According to the meta-analysis, this process is subserved by the mid section of the left middle temporal gyrus. The chronometric data suggest a time window between 175 and 250 ms in picture naming. Lexical selection in word production was targeted in a recent MEG study on picture naming (Maess, Friederici, Damian, Meyer, & Levelt, 2002) by use of a semantic category interference paradigm. In this paradigm, the naming of objects in blocks comprising other objects of the same semantic category is slowed down compared to the naming of objects in semantically heterogeneous blocks. This effect is due to competition between similar lexical entries. For subjects showing the behavioural effect, Maess et al. found significant activation differences between the same-category and the different-category conditions in the mid section of the left middle temporal lobe in the time window of 150–225 ms post-stimulus. These data are in line with the assumed neural correlate and time window for lemma selection.

3.1.2. Lexical phonological code retrieval

According to the meta-analysis, this process is subserved by the posterior sections of the middle and the superior temporal gyrus of the left hemisphere. The chronometric data suggest a time window between 250 and 330 ms in picture naming. This time window is in agreement with the MEG studies of Salmelin et al. (1994), reporting posterior middle temporal gyrus activation from 200 to 400 ms, and Levelt et al. (1998), reporting posterior superior temporal gyrus activation in the time window 275–400 ms (see Appendix D).

3.1.3. Syllabification

According to the meta-analysis, this process is subserved by the left posterior inferior frontal gyrus. The chronometric data suggest a time window between 330 and 455 ms in picture naming. The MEG data of Salmelin et al. (1994) showing activation of Broca's area between 400 and 600 ms are compatible with this time window.

3.1.4. Self-monitoring

In the meta-analysis, we identified bilateral superior temporal areas as possible neural correlates of self-monitoring. Assuming that the internal loop of self-monitoring takes the output of syllabification as input, the earliest activation of these areas may be expected immediately after the first spelled-out segment is used for the production of a phonological word, that is after 355 ms (see Section 1.2). The time windows observed in two MEG studies (275–400 ms for the left mid and posterior superior temporal gyrus, 300–800 ms for the right mid superior temporal gyrus; Levelt et al., 1998; Salmelin et al., 1994) are compatible with a role of the superior temporal regions in internal self-monitoring.

In sum, the available timing data are compatible with the functional roles that were assigned to the different cortical areas on the basis of the meta-analysis. Although the observed time windows do not in all cases correspond exactly to the intervals predicted on the basis of the chronometric data, they cover at least a major part of the expected intervals, and they show the expected temporal order. It should be kept in mind, however, that the number of MEG studies providing these data is still very small. Fig. 5 summarizes the timing information for the regions related to the core processes of word production.

3.2. Time windows of interaction between word production and perception

Word production is affected by simultaneous word perception processes. In particular, there are two robust effects known as semantic interference and phonological facilitation. Semantic interference means that picture naming is slowed down when subjects are visually or auditorily presented with semantically related distracters. It is assumed that the activated distracter lemma interferes with lemma selection in the production pathway (see arrow 1 in Fig. 2). Phonological facilitation (or priming) means that subjects can name a picture faster when presented with a phonologically similar distracter than when presented with a dissimilar distracter. This effect is assumed to occur through pre-activation of lexical phonological codes or spelled-out segments (see arrows 2 and 3 in Fig. 2). The chronometric data presented in Section 1.2 provide some constraints for the time windows during which the two kinds of interactions occur. We will now use this information to examine whether it is possible to identify regions activated during visual or auditory word perception, which might feed the word production pathway with interfering or facilitating input.

3.2.1. Semantic interference

As calculated in Section 1.2, lemma activation in word listening should occur between 325 and 400 ms post word onset, since interference is strongest when distracter words are presented about 150 ms earlier than the picture to be named. The data of Simos, Breier, Fletcher, Bergman, and Papanicolaou (2000) suggest compatible time windows for all areas involved in word listening (Appendix D), except for the right mid superior temporal gyrus, which seems to be activated too early (100–300 ms), and the left mid superior temporal gyrus, which seems to be activated too late (400–600 ms).

Given that semantically related written word distracters show their maximal effect when presented simultaneously with the picture to be named, their lemma activation should peak at the time of lemma selection in picture naming, that is 175–250 ms. Concentrating first on areas for which the reported time windows overlap between studies, temporal and frontal regions are activated too late (300–400 ms). Regions showing at least partially compatible timing are the left angular gyrus (200–400 ms) and the cuneus (150–200 ms). Both, however, have not been reliably found in PET and fMRI studies of reading. The earliest activations (200 ms) observed by Salmelin et al. (Salmelin et al., 2000; Salmelin, Service, Kiesilä, Uutela, & Salonen, 1996) suggest that depending on the experimental conditions left posterior temporal areas and the left posterior inferior frontal gyrus may be activated in time to play a role in lemma activation following visual word presentation. The same holds for medial and lateral occipital areas (150–200 ms).

In sum, the timing information available from word reading and listening studies does not unambiguously identify regions subserving visually or phonologically driven lemma access in the right time window to interfere with word production.

3.2.2. *Phonological facilitation*

Given that typical SOAs for phonological facilitation by spoken word distracters are 0–150 ms, the activation of the phonological code in word listening should occur between 100 and 330 ms. The data provided by Simos et al. (2000) suggest that only right mid and posterior temporal areas are activated in time, whereas left temporal areas are activated somewhat later (> 300 ms). Note, however, that Simos et al. showed earlier activation of left temporal areas in dyslexic subjects. The rather late activation of left temporal areas observed by Simos et al. seems, furthermore, difficult to reconcile with the MEG evidence for phonological processing in the left mid superior temporal gyrus in a time window of 150–250 ms provided by Phillips et al. (2000). Clearly more studies on the time course of cortical activations during word listening are needed to come to firm conclusions about the role of the left temporal lobe in phonological word form retrieval. At present, the available timing data for word listening do not allow the identification of a source area from which information resulting in phonological priming might be relayed.

SOAs for phonological priming by written word distracters may range considerably from 200 ms before to 200 ms after picture onset. Taking into account the time window of phonological code retrieval (250–330 ms), this gives us an estimated range of 50–530 ms for the activation of lexical or sublexical phonological representations (neglecting the unknown duration of the grapheme-to-phoneme conversion process). The MEG experiments on word reading agree that during this time window both the left posterior inferior frontal gyrus and several temporal regions (bilateral mid superior temporal gyri, left posterior middle and superior temporal gyri) are activated. In addition, the MEG studies suggest bilateral posterior parietal activations in this time window. According to the meta-analysis results, however, parietal activations have at present not been reliably found in hemodynamic studies of word reading.

Broca's area or adjacent neural tissue have been suggested to be involved in sublexical grapheme-to-phoneme conversion (Fiez et al., 1999; Hagoort et al., 1999; Pugh et al., 1996). Since the output of this process has the right format to interact with the post-lexical segmental representation in word production, it is plausible that Broca's area should be involved in phonological priming from written words or pseudowords.

For left posterior temporal regions, Pugh et al. (1996) and Fiez et al. (1999) report lexicality effects with stronger activations for words than for pseudowords. These findings are in line with the results of the present meta-analysis (see Section 2.5 and Appendix B). A phonological priming effect involving left posterior temporal areas, therefore, might be based on the pre-activation of *lexical* phonological codes. Note, however, that according to our meta-analysis the posterior section of the left superior temporal gyrus may also be involved in internal self-monitoring (together with the bilateral mid sections of the superior temporal gyri), although no estimate for the time window of self-monitoring in reading is available.

3.3. Integrating temporal and spatial information – a tentative flow-chart of word production

Taking into account all three sources of information, i.e. the spatial distribution of activations, the time course of activations in picture naming, and the chronometric data, a tentative but consistent picture of the flow of activation in word production has emerged. It should be kept in mind that the exact time windows given here are based on picture naming data, and may be shifted in time for other word production tasks.

Visual and conceptual lead-in processes involving occipital and ventrotemporal regions converge within 175 ms from stimulus onset on a lexical concept to be expressed. The best-fitting lexical item is selected within the following 75 ms. This conceptually driven lexical selection seems to engage the mid section of the left middle temporal gyrus in the time window between 150 and 225 ms. During this time window semantically related written or spoken distracter words can interfere with lexical selection, but it is at present unclear from which cortical areas involved in processing the distracters the interfering information is relayed.

From about 200 ms onwards, the information spreads to the posterior temporal lobe (Wernicke's area), where the lexically stored phonological code of the word is retrieved in the time window between 250 and 330 ms. During this time window the lexical phonological codes may be primed by phonologically related written word distracters, which activate Wernicke's area from about 240–260 ms onwards, when presented at typical SOAs of 40–60 ms before picture onset. In the same time window, written lexical as well as non-lexical (pseudoword) distracters activate Broca's area. Given the role of Broca's area in grapheme-to-phoneme conversion, it is plausible that the sublexical phonological representations resulting from this process pre-activate the spelled-out segments of the production pathway, which are relayed anteriorly to Broca's area for post-lexical syllabification. Syllabification results in a phonological word at about 455 ms.

Within another 145 ms the phonological word is phonetically encoded (with possible contributions of SMA and cerebellum) and sensorimotor areas involved in articulation become active. Internal self-monitoring, probably involving bilateral superior temporal areas, begins as soon as the production of the phonological word starts with the first segment (at about 355 ms). After articulation has begun, the same areas are involved in listening to one's own speech (external self-monitoring).

4. Conclusions

The spoken language user constantly operates a dual system, perceiving and producing utterances. These systems not only alternate, but in many cases they partially or wholly operate in concert. Levelt (2001) reviewed some of the experimental evidence demonstrating effects on utterance production that are caused by concomitant perceptual input. The two systems not only operate in concert, but they interact in specific ways at different levels of processing. These facts, it was argued, can be used

to guide the search for what is shared among the neural networks that support speech perception and production. The present paper practices that approach in a comprehensive meta-analysis of the relevant imaging literature on word production (82 experiments) and on auditory word/non-word perception (26 experiments). A componential analysis of the tasks involved (picture naming, word generation, word and pseudoword reading, word and pseudoword listening) provided the core operations involved in each of the tasks. All tasks involving a particular core operation (such as phonological code retrieval) should share activation of the relevant brain region(s). In addition to taking spatial overlap of component perceptual and production regions into account, we also analyzed whatever is known about the time course of activations. The latter is a potentially powerful addition to an earlier meta-analysis (Indefrey & Levelt, 2000), given the assumption that perceptual effects on production require temporal overlap of activations in the relevant region(s).

The present analysis supports and extends our earlier findings, specifying regions and time windows of activation for the core processes of word production: lexical selection, phonological code retrieval, syllabification and phonetic/articulatory preparation. It is, in particular, satisfying to observe that the time course of these activations is, on the whole, compatible with the temporal constraints that processes of auditory word and non-word perception impose on the component production processes they affect.

Still, this is only a beginning. We are in need of drastically more precise measurements of the chronometric regime within the production and perceptual networks that support the production and perception of speech and their interactions. Only theory-driven ERP and MEG studies, such as those reviewed in Section 3 of the present paper, can provide that critical information.

Appendix A. Reliably activated and not activated cerebral regions based on 82 word production experiments

The number of times a region was reported as activated (found) is given in proportion to the number of experiments in which this region was within the field of view (covered). For a mean number of 12.4 reported areas, the chance probability was smaller than 0.1 for three or less reports out of 60 or more experiments (region reliably not activated) and 11 or more reports out of 57–63 experiments (12 or more out of 64–70, 13 or more out of 71–73, region reliably activated). The abbreviations of gyri and subcortical structures follow Talairach and Tournoux (1988) except for SMA, supplementary motor area. Cu, Cuneus; Ga, angular gyrus; GF, fusiform gyrus; GFd, medial frontal gyrus; GFs, GFm, GFi, superior, middle, and inferior frontal gyrus; GH, parahippocampal gyrus; GL, lingual gyrus; GO, orbital gyri; Go, occipital gyri; GPoC, postcentral gyrus; GPrC, precentral gyrus; GR, gyrus rectus; Gs, gyrus subcallosus; Gsm, supramarginal gyrus; GTs, GTm, GTi, superior, middle, and inferior temporal gyrus; LPs, LPi, superior and inferior parietal lobule; NL, lenticular nucleus; PCu, precuneus; Sca, calcarine sulcus.

All production tasks

No. experiments
overt, covert

82
35, 47

mean no. subjects
mean no. regions

11.8
12.4

reliably activated
reliably not activated

			found	covered
Frontal	R anterior	GfS	0	64
		GfM	5	64
		GfI	1	64
		GfD	2	64
		GR, GO	3	65
	posterior	GfS	1	61
		GfM	9	64
		GfI	16	66
		GfD	3	64
		GR, Gsc	1	64
	motor	ventGPrc	18	68
		dorsGPrc	14	68
		SMA	18	66
	L anterior	GfS	5	64
		GfM	12	65
		GfI	7	65
		GfD	4	64
		GR, GO	6	67
	posterior	GfS	4	62
		GfM	22	66
		GfI	51	71
		GfD	13	64
		GR, Gsc	1	64
	motor	ventGPrc	32	71
		dorsGPrc	21	69
		SMA	26	66
Temporal	R anterior	GtS	5	65
		GtM	0	64
		GH	0	64
	mid	GtS	23	67
		GtM	16	65
		GtI	1	62
	posterior	GF	3	66
		GH	2	67
		GL	10	65
	L anterior	GtS	13	71
		GtM	0	71
		GH	0	68
	mid	GtS	30	73
		GtM	24	71
		GtI	8	70
	posterior	GF	9	68
		GH	5	71
		GL	29	72
Insula	R anterior	GtS	18	65
		GL	3	63
	L anterior	GtS	30	69
		GL	5	64

			found	covered
Parietal	R sensory	ventGPoc	5	64
		dorsGPoc	1	64
		LPI	1	64
		LPS	0	64
		Lpc	1	64
	posterior	PCu	0	64
		LPS	4	64
		Lpi	2	64
		Gsm	2	64
		Ga	2	64
	L sensory	ventGPoc	8	67
		dorsGPoc	1	66
		LPI	6	67
		LPS	0	65
		Lpc	0	64
Occipital	R medial	PCu	0	64
		LPS	4	65
		LPI	2	66
		Gsm	5	67
		Ga	4	66
	L medial	PCu	1	64
		Sca	5	64
		Cu	9	64
		GL	18	68
		GF	14	66
Cingulum	R anterior	Go	15	66
		Sca	8	64
		Cu	9	66
		GL	22	68
		GF	16	66
Sub-cortical structures	R Caudate	Go	19	66
		Sca	8	64
		Cu	9	66
		GL	22	68
		GF	16	66
Cerebellum	R anterior	Go	19	66
		Sca	8	64
		Cu	9	66
		GL	22	68
		GF	16	66

Appendix B. Reliably activated cerebral regions for all tasks

Note that depending on the mean number of regions reported per experiment the relative number of reports required to pass the reliability criterion may vary between tasks. For abbreviations of gyri and subcortical structures see Appendix A.

Task		picture naming	word generation	word reading	pseudoword reading	word listening	pseudoword listening
No. experiments		23	25	24	10	20	6
overt, covert		19, 4	1, 24	10, 14	5, 5		
mean no. subjects		20.2	6.1	10.2	10.3	6.7	9.8
mean no. regions		10.5	12.6	14.6	12	6	5

reliably activated																	
		found	covered	found	covered	found	covered	found	covered	found	covered	found	covered	found	covered	found	covered
Frontal	R anterior	GFs	0	11	0	23	0	22	0	8	0	15	0	5			
		GFm	0	11	4	23	1	22	0	8	0	16	0	5			
		GFI	0	11	0	23	1	22	0	8	1	15	0	5			
		GFd	1	11	0	23	1	22	0	8	0	15	0	4			
		GR_GO	2	12	1	23	0	22	0	8	0	14	0	4			
	posterior	GFs	0	11	0	20	1	22	0	8	0	16	0	5			
		GFm	0	11	6	23	2	22	1	8	0	15	0	5			
		GFI	4	11	4	23	5	23	3	9	1	16	0	5			
		GFd	1	11	1	23	1	22	0	8	0	15	0	4			
		GR_Gsc	1	11	0	23	0	22	0	8	0	15	0	4			
	motor	ventGPrc	4	12	2	24	8	23	4	9	1	15	0	5			
		dorsGPrc	2	12	1	24	8	23	3	9	1	15	0	5			
		SMA	3	12	6	24	7	22	2	8	0	16	0	5			
	L anterior	GFs	0	11	5	23	0	22	0	8	0	15	0	5			
		GFm	1	11	7	24	3	22	1	8	0	16	0	5			
		GFI	0	11	4	24	1	22	2	8	1	15	0	5			
		GFd	0	11	3	23	1	22	0	8	0	15	0	4			
		GR_GO	2	12	1	23	2	23	1	9	0	14	0	4			
	posterior	GFs	1	12	2	20	1	22	0	8	0	16	0	5			
		GFm	2	13	12	23	7	22	1	8	1	15	0	5			
		GFI	10	16	23	24	13	23	5	8	4	17	0	5			
		GFd	0	11	11	23	2	22	0	8	1	15	0	4			
		GR_Gsc	1	11	0	23	0	22	0	8	0	15	0	4			
	motor	ventGPrc	8	14	7	24	10	23	7	10	1	15	0	5			
		dorsGPrc	3	13	7	24	9	23	2	9	1	15	0	5			
		SMA	3	12	13	24	7	22	3	8	0	16	0	5			
Temporal	R anterior	GTs	1	11	1	22	2	23	1	9	5	16	3	5			
		GTm	0	11	0	23	0	22	0	8	4	15	2	5			
		GH	0	11	0	23	0	22	0	8	2	14	0	4			
	mid	GTs	4	12	8	23	7	23	4	9	14	17	4	5			
		GTm	2	12	10	23	4	22	0	8	7	16	2	5			
		GTi	1	11	0	21	0	22	0	8	0	15	0	5			
		GF	1	12	1	24	0	22	1	8	0	14	0	4			
		GH	2	12	0	23	0	23	0	9	0	15	0	5			
	posterior	GTs	2	11	2	22	5	23	1	9	5	15	2	5			
		GTm	3	11	3	22	2	22	0	8	1	16	1	5			
		GTi	1	13	1	22	1	22	1	8	0	15	0	5			
		GF	6	13	2	24	3	23	2	9	0	14	0	4			
		GL	0	11	0	23	1	22	1	9	0	14	0	4			
	L anterior	GTs	5	18	3	22	5	23	0	8	7	18	3	5			
		GTm	0	18	0	23	0	22	0	8	3	17	2	5			
		GH	0	15	0	23	0	22	0	8	1	15	0	4			
	mid	GTs	8	19	6	22	11	23	5	9	17	19	4	5			
		GTm	7	19	12	22	3	22	2	8	8	18	2	5			
		GTi	6	19	0	21	2	22	0	8	0	17	0	5			
		GF	5	15	0	23	4	22	0	8	0	15	0	4			
		GH	2	16	0	23	2	23	1	9	0	15	0	4			
	posterior	GTs	8	18	10	22	9	23	2	9	11	18	3	6			
		GTm	7	18	14	22	7	22	1	8	5	18	2	5			
		GTi	7	19	2	21	5	22	1	8	2	17	0	5			
		GF	10	16	7	24	11	23	3	9	0	15	0	4			
		GL	0	14	0	23	2	22	1	8	0	16	0	5			

Task			picture naming		word generation		word reading		pseudoword reading		word listening		pseudoword listening		
			found	covered	found	covered	found	covered	found	covered	found	covered	found	covered	
Parietal	R sensory	ventGPoc	1	11	0	23	3	22	1	8	1	15	0	5	
		dorsGPoc	0	11	0	23	1	22	0	8	1	15	0	5	
		anterior	LPI	0	11	0	23	1	22	0	8	0	15	0	5
			LPs	0	11	0	23	0	22	0	8	0	15	0	5
			Lpc	0	11	0	23	1	22	0	8	0	14	0	4
	PCu		0	11	0	23	0	22	0	8	0	14	0	4	
	posterior	LPs	0	11	0	23	2	22	2	8	0	15	0	5	
		Lpi	0	11	0	23	2	22	0	8	0	15	0	5	
		Gsm	1	11	0	23	1	22	0	8	0	15	0	5	
		Ga	1	11	0	23	1	22	0	8	0	15	0	5	
		PCu	0	11	0	23	1	22	0	7	0	14	0	4	
	L sensory	ventGPoc	3	14	0	23	4	22	1	7	1	15	0	5	
		dorsGPoc	0	13	0	23	1	22	0	7	1	15	0	5	
		anterior	LPI	3	14	0	23	3	22	0	7	0	16	0	5
			LPs	0	12	0	23	0	22	0	7	0	16	0	5
			Lpc	0	11	0	23	0	22	0	7	0	14	0	4
			PCu	0	11	0	23	0	22	0	7	0	14	0	4
		posterior	LPs	0	12	1	23	2	22	1	7	0	16	0	5
			LPI	0	13	0	23	2	22	0	7	0	16	0	5
			Gsm	3	14	1	23	1	22	0	7	1	17	0	5
			Ga	1	13	1	23	2	22	0	7	0	16	0	5
PCu		0	11	0	23	1	22	0	7	0	14	0	4		
Occipital	R medial	Sca	2	11	0	23	3	22	0	7	0	15	0	4	
		Cu	2	11	0	23	4	22	2	7	0	14	0	4	
		GL	5	12	1	24	9	23	3	8	1	15	0	4	
		GF	4	11	0	23	7	23	3	8	0	14	0	4	
	lateral	Go	5	12	2	24	6	22	2	7	0	16	0	5	
		Sca	3	11	1	23	4	22	0	7	0	15	0	4	
	L medial	Cu	2	11	0	23	5	23	1	8	0	14	0	4	
		GL	7	12	1	24	10	23	4	8	1	15	0	4	
		GF	4	11	0	23	7	23	5	8	0	14	0	4	
		Go	7	12	1	24	9	22	2	7	0	16	0	5	
	lateral	Sca	1	11	3	23	1	22	0	7	0	14	0	3	
		mid	4	11	2	23	0	22	1	7	0	14	0	3	
		posterior	0	11	0	23	1	22	0	7	0	14	0	3	
L anterior		1	11	6	23	3	22	0	7	1	14	0	3		
Insula	R	anterior	2	12	10	23	5	22	1	7	0	14	0	3	
		posterior	0	11	2	22	1	22	0	7	0	14	0	3	
		L anterior	6	13	10	23	12	24	2	8	0	14	0	3	
	L	posterior	0	11	3	22	2	23	0	7	0	14	0	3	
		Caudate	2	11	2	24	0	20	0	7	1	14	0	3	
	Sub-cortical structures	NL	1	9	2	23	2	20	0	7	0	15	0	3	
		Thalamus	2	10	7	24	5	20	0	7	0	15	0	3	
		Hypothalamus	0	9	0	23	0	20	0	7	0	14	0	3	
		Midbrain	3	10	1	24	1	20	0	7	1	14	0	3	
		Hippocampus	0	9	0	23	0	20	0	7	0	14	0	3	
		Caudate	1	9	4	23	1	20	0	7	0	14	0	3	
		NL	2	9	8	23	5	20	0	7	0	15	0	3	
	Cerebellum	R	Thalamus	5	11	11	24	6	21	2	8	0	15	0	3
Hypothalamus			0	9	0	23	0	20	0	7	0	14	0	3	
Midbrain			2	10	1	24	2	20	0	7	0	14	0	3	
Hippocampus			0	9	0	23	2	20	0	7	0	14	0	3	
L		medial	6	9	8	19	9	21	4	8	1	15	0	3	
		lateral	3	9	8	18	5	20	3	7	1	15	0	3	
		lateral	5	9	8	19	4	21	3	8	1	15	0	3	

Appendix C. Reliably activated cerebral regions for all word production experiments with overt responses and silent control conditions and all word production experiments with covert responses

Column 'found' gives the number of times a region was reported as activated. Column 'covered' gives the number of experiments in which this region was within the field of

view. Column ‘proportion’ gives the relative frequency of reports on this region. For abbreviations of gyri and subcortical structures see Appendix A.

			overt production with silent control			covert production		
			found	covered	proportion	found	covered	proportion
Parietal	R sensory	ventGPoc	4	13	0.31	1	43	0.02
		dorsGPoc	1	13	0.08	0	43	0.00
	anterior	LPi	0	13	0.00	1	43	0.02
		LPs	0	13	0.00	0	43	0.00
		Lpc	0	13	0.00	1	43	0.02
		PCu	0	13	0.00	0	43	0.00
	posterior	LPs	1	13	0.08	3	43	0.07
		Lpi	1	13	0.08	1	43	0.02
		Gsm	2	13	0.15	0	43	0.00
		Ga	2	13	0.15	0	43	0.00
		PCu	0	13	0.00	1	43	0.02
	L sensory	ventGPoc	6	16	0.38	2	43	0.05
		dorsGPoc	1	15	0.07	0	43	0.00
		LPi	3	16	0.19	2	43	0.05
		LPs	0	14	0.00	0	43	0.00
		Lpc	0	13	0.00	0	43	0.00
		PCu	0	13	0.00	0	43	0.00
		LPs	0	14	0.00	4	43	0.09
		Lpi	1	15	0.07	1	43	0.02
		Gsm	4	16	0.25	1	43	0.02
		Ga	3	15	0.20	1	43	0.02
Occipital	R medial	Sca	3	13	0.23	2	43	0.05
		Cu	4	13	0.31	5	43	0.12
		GL	11	17	0.65	7	43	0.16
		GF	6	15	0.40	8	43	0.19
		Go	5	15	0.33	9	43	0.21
	L medial	Sca	3	13	0.23	4	43	0.09
		Cu	6	15	0.40	3	43	0.07
		GL	12	17	0.71	7	43	0.16
		GF	9	15	0.60	7	43	0.16
		Go	6	15	0.40	11	43	0.26
	lateral							
Cingulum	R anterior		1	13	0.08	3	43	0.07
		mid	3	13	0.23	4	43	0.09
		posterior	0	13	0.00	1	43	0.02
	L anterior		2	13	0.15	7	43	0.16
		mid	4	13	0.31	6	43	0.14
		posterior	0	13	0.00	2	43	0.05
Sub-cortical structures	R	Caudate	3	13	0.23	1	42	0.02
		NL	0	10	0.00	4	42	0.10
		Thalamus	6	12	0.50	8	42	0.19
		Hypothalamus	0	10	0.00	0	42	0.00
		Midbrain	4	12	0.33	1	42	0.02
		Hippocampus	0	10	0.00	0	42	0.00
	L	Caudate	1	10	0.10	4	42	0.10
		NL	2	10	0.20	11	42	0.26
		Thalamus	10	15	0.67	14	42	0.33
		Hypothalamus	0	10	0.00	0	42	0.00
		Midbrain	3	12	0.25	2	42	0.05
		Hippocampus	1	10	0.10	0	42	0.00
Cerebellum	R	medial	11	12	0.92	13	38	0.34
		lateral	4	10	0.40	13	37	0.35
	L	medial	9	12	0.75	10	38	0.26
		lateral	8	12	0.67	8	37	0.22

		overt production with silent control			covert production			
No. experiments		26			47			
mean no. subjects		18.5			8.3			
mean no. regions		15.3			11.8			
reliably activated								
		found covered proportion			found covered proportion			
Frontal	R anterior	GFs	0	13	0.00	0	43	0.00
		GFm	0	13	0.00	5	43	0.12
		GFi	0	13	0.00	1	43	0.02
		GFd	1	13	0.08	1	43	0.02
		GR_GO	3	15	0.20	0	42	0.00
	posterior	GFs	0	13	0.00	1	40	0.03
		GFm	0	13	0.00	9	43	0.21
		GFi	7	15	0.47	8	43	0.19
		GFd	1	13	0.08	2	43	0.05
		GR_Gsc	1	13	0.08	0	43	0.00
	motor	ventGPrC	15	17	0.88	3	43	0.07
		dorsGPrC	7	17	0.41	7	43	0.16
		SMA	8	15	0.53	10	43	0.23
	L anterior	GFs	0	13	0.00	5	43	0.12
		GFm	1	13	0.08	11	44	0.25
		GFi	0	13	0.00	7	44	0.16
		GFd	0	13	0.00	4	43	0.09
		GR_GO	5	17	0.29	1	42	0.02
	posterior	GFs	1	14	0.07	2	40	0.05
		GFm	1	15	0.07	19	43	0.44
		GFi	11	17	0.65	35	46	0.76
		GFd	0	13	0.00	13	43	0.30
		GR_Gsc	1	13	0.08	0	43	0.00
	motor	ventGPrC	17	19	0.89	13	44	0.30
		dorsGPrC	8	18	0.44	13	43	0.30
		SMA	9	15	0.60	17	43	0.40
Temporal	R anterior	GTs	3	15	0.20	1	42	0.02
		GTm	0	13	0.00	0	43	0.00
		GH	0	13	0.00	0	43	0.00
	mid	GTs	14	17	0.82	9	42	0.21
		GTm	5	15	0.33	11	42	0.26
		GTi	1	13	0.08	0	41	0.00
		GF	2	15	0.13	1	43	0.02
		GH	0	13	0.00	1	46	0.02
	posterior	GTs	5	15	0.33	5	42	0.12
		GTm	4	13	0.31	4	42	0.10
		GTi	2	16	0.13	2	41	0.05
		GF	7	18	0.39	5	43	0.12
		GL	1	14	0.07	1	43	0.02
	L anterior	GTs	6	20	0.30	5	42	0.12
		GTm	0	19	0.00	0	43	0.00
		GH	0	16	0.00	0	43	0.00
	mid	GTs	17	22	0.77	10	42	0.24
		GTm	8	20	0.40	15	42	0.36
		GTi	5	20	0.25	2	41	0.05
		GF	3	16	0.19	2	43	0.05
		GH	0	16	0.00	4	46	0.09
	posterior	GTs	13	21	0.62	13	42	0.31
		GTm	7	19	0.37	20	42	0.48
		GTi	3	20	0.15	8	41	0.20
		GF	13	20	0.65	15	43	0.35
		GL	0	15	0.00	3	43	0.07
Insula	R anterior	4	15	0.27	13	42	0.31	
	posterior	1	13	0.08	2	42	0.05	
	L anterior	10	17	0.59	19	44	0.43	
	posterior	2	14	0.14	3	42	0.07	

References

- Abdullaev, Y. G., & Melnichuk, K. V. (1995). *Neuronal activity of human caudate nucleus in cognitive tasks* (Tech. Rep. No. 95-09). Eugene: University of Oregon.
- Abdullaev, Y. G., & Posner, M. I. (1997). Time course of activating brain areas in generating verbal associations. *Psychological Science*, 8, 56–59.
- Bachoud-Lévi, A.-C., Dupoux, E., Cohen, L., & Mehler, J. (1998). Where is the length effect? A cross-linguistic study of speech production. *Journal of Memory and Language*, 39, 331–346.
- Beauregard, M., Chertkow, H., Bub, D., Murtha, S., Dixon, R., & Evans, A. (1997). The neural substrate for concrete, abstract, and emotional word lexica: a positron emission tomography study. *Journal of Cognitive Neuroscience*, 9, 441–461.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Springer, J. A., Kaufman, J. N., & Possing, E. T. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cerebral Cortex*, 10, 512–528.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Rao, S. M., & Cox, R. W. (1996). Function of the left planum temporale in auditory and linguistic processing. *Brain*, 119, 1239–1247.
- Binder, J. R., Rao, S. M., Hammeke, T. A., Yetkin, F. Z., Jesmanowicz, A., Bandettini, P. A., Wong, E. C., Estkowski, L. D., Goldstein, M. D., Haughton, V. M., & Hyde, J. S. (1994). Functional magnetic resonance imaging of human auditory cortex. *Annals of Neurology*, 35, 662–672.
- Bookheimer, S. Y., Zeffiro, T. A., Blaxton, T., Gaillard, W., & Theodore, W. (1995). Regional cerebral blood flow during object naming and word reading. *Human Brain Mapping*, 3, 93–106.
- Brunswick, N., McCrory, E., Price, C. J., Frith, C. D., & Frith, U. (1999). Explicit and implicit processing of words and pseudowords by adult developmental dyslexics. A search for Wernicke's Wortschatz? *Brain*, 122, 1901–1917.
- Buckner, R. L., Raichle, M. E., & Petersen, S. E. (1995). Dissociation of human prefrontal cortical areas across different speech production tasks and gender groups. *Journal of Neurophysiology*, 74(5), 2163–2173.
- Butterworth, B. (1989). Lexical access in speech production. In W. Marslen-Wilson (Ed.), *Lexical representation and process*. Cambridge, MA: MIT Press.
- Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology*, 14, 177–208.
- Cardebat, D., Démonet, J. F., Puel, M., Agniel, A., Viallard, G., & Celsis, P. (1998). Brain correlates of memory processes in patients with dementia of Alzheimer's type – a SPECT activation study. *Journal of Cerebral Blood Flow and Metabolism*, 18, 457–462.
- Chao, L. L., & Martin, A. (1999). Cortical regions associated with perceiving, naming, and knowing about colors. *Journal of Cognitive Neuroscience*, 11, 25–35.
- Clark, E. V. (1997). Conceptual perspective and lexical choice in acquisition. *Cognition*, 64, 1–7.
- Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Hénaff, M.-A., & Michel, F. (2000). The visual word form area. Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain*, 123, 291–307.
- Crivello, F., Tzourio, N., Poline, J. B., Woods, R. P., Mazziotta, J. C., & Mazoyer, B. (1995). Intersubject variability in functional neuroanatomy of silent verb generation: assessment by a new activation detection algorithm based on amplitude and size information. *NeuroImage*, 2, 253–263.
- Cutler, A., & Clifton, C. (1999). Comprehending spoken language: a blueprint of the listener. In C. M. Brown, & P. Hagoort (Eds.), *The neurocognition of language*. Oxford: Oxford University Press.
- Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D., & Damasio, A. R. (1996). A neural basis for lexical retrieval. *Nature*, 380, 499–505.
- Damian, M. F., & Martin, R. (1999). Semantic and phonological cues interact in single word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 345–361.
- Damian, M. F., Vigliocco, G., & Levelt, W. J. M. (2001). Effects of semantic context in the naming of pictures and words. *Cognition*, 81, B77–B86.
- David, A. S., & Busatto, G. (1999). Die Halluzination. Eine Störung von Gehirn und Geist. *Nervenheilkunde*, 18, 104–115.

- David, A. S., Woodruff, P. W. R., Howard, R., Mellers, J. D. C., Brammer, M., Bullmore, E., Wright, I., Andrew, C., & Williams, S. C. R. (1996). Auditory hallucinations inhibit exogenous activation of auditory association cortex. *NeuroReport*, 7, 932–936.
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 93, 283–321.
- Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, 104, 801–838.
- Démonet, J. F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J. L., Wise, R., Rascol, A., & Frackowiak, R. (1992). The anatomy of phonological and semantic processing in normal subjects. *Brain*, 115, 1753–1768.
- Dhankhar, A., Wexler, B. E., Fulbright, R. K., Halwes, T., Blamire, A. M., & Shulman, R. G. (1997). Functional magnetic resonance imaging assessment of the human brain auditory cortex response to increasing word presentation rates. *Journal of Neurophysiology*, 77, 476–483.
- Dierks, T., Linden, D. E. J., Jandl, M., Formisano, E., Goebel, R., Lanfermann, H., & Singer, W. (1999). Activation of Heschl's gyrus during auditory hallucinations. *Neuron*, 22, 615–621.
- Dronkers, M. F. (1996). A new brain region for coordinating speech articulation. *Nature*, 384, 159–161.
- Etard, O., Mellet, E., Papathanassiou, D., Benali, K., Houde, O., Mazoyer, B., & Tzourio-Mazoyer, N. (2000). Picture naming without Broca's and Wernicke's area. *NeuroReport*, 11, 617–622.
- Eulitz, C., Elbert, T., Bartenstein, P., Weiller, C., Muller, S. P., & Pantev, C. (1994). Comparison of magnetic and metabolic brain activity during a verb generation task. *NeuroReport*, 6, 97–100.
- Ferrand, L., Segui, J., & Grainger, J. (1996). Masked priming of word and picture naming: the role of syllabic units. *Journal of Memory and Language*, 35, 708–723.
- Fiez, J., Petersen, S., & Raichle, M. (1996). Identification of two pathways used for verbal response selection. In C. Chase, G. Rosen, & G. Sherman (Eds.), *Developmental dyslexia: neural, cognitive, and genetic mechanisms*. Timonium, MD: York Press.
- Fiez, J. A., Balota, D. A., Raichle, M. E., & Petersen, S. E. (1999). Effects of lexicality, frequency and spelling-to-sound consistency on the functional anatomy of reading. *Neuron*, 24, 205–218.
- Fiez, J. A., Raichle, M. E., Balota, D. A., Tallal, P., & Petersen, S. E. (1996). PET activation of posterior temporal regions during auditory word presentation and verb generation. *Cerebral Cortex*, 6, 1–10.
- Fiez, J. A., Raichle, M. E., Miezin, F. M., Petersen, S. E., Tallal, P., & Katz, W. F. (1995). PET studies of auditory and phonological processing – effects of stimulus characteristics and task demands. *Journal of Cognitive Neuroscience*, 7, 357–375.
- Friston, K. J., Price, C. J., Fletcher, P., Moore, C., Frackowiak, R. S. J., & Dolan, R. J. (1996). The trouble with cognitive subtraction. *NeuroImage*, 4, 97–104.
- Garrett, M. F. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Speech and talk (I). Language production*, New York: Academic Press.
- Glaser, W. R., & Döngelhoff, F.-J. (1984). The time-course of picture-word interference. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 640–654.
- Haglund, M. M., Berger, M. S., Shamseldin, M., Lettich, E., & Ojemann, G. A. (1994). Cortical localization of temporal lobe language sites in patients with gliomas. *Neurosurgery*, 34, 567–576.
- Hagoort, P., Indefrey, P., Brown, C., Herzog, H., Steinmetz, H., & Seitz, R. J. (1999). The neural circuitry involved in the reading of German words and pseudowords: a PET study. *Journal of Cognitive Neuroscience*, 11, 383–398.
- Hamberger, M. J., Goodman, R. R., Perrine, K., & Tamny, T. (2001). Anatomic dissociation of auditory and visual naming in the lateral temporal cortex. *Neurology*, 56, 56–61.
- Herbster, A. N., Mintun, M. A., Nebes, R. D., & Becker, J. T. (1997). Regional cerebral blood flow during word and nonword reading. *Human Brain Mapping*, 5, 84–92.
- Hirano, S., Kojima, H., Naito, Y., Honjo, I., Kamoto, Y., Okazawa, H., Ishizu, K., Yonekura, Y., Nagahama, Y., Fukuyama, H., & Konishi, J. (1997). Cortical processing mechanism for vocalization with auditory verbal feedback. *NeuroReport*, 8, 2379–2382.
- Howard, D., Patterson, K., Wise, R., Brown, W. D., Friston, K., Weiller, C., & Frackowiak, R. (1992). The cortical localization of the lexicons. Positron emission tomography evidence. *Brain*, 115, 1769–1782.

- Indefrey, P. (1997). PET research in language production. In W. Hulstijn, H. E. M. Peters, & P. H. H. M. Van Lieshout (Eds.), *Speech production: motor control, brain research and fluency disorders*. Amsterdam: Elsevier.
- Indefrey, P., & Levelt, W. J. M. (2000). The neural correlates of language production. In M. Gazzaniga (Ed.), *The new cognitive neurosciences*. Cambridge, MA: MIT Press.
- Indefrey, P., Hagoort, P., Herzog, H., Seitz, R. J., & Brown, C. M. (2001). Syntactic processing in left prefrontal cortex is independent of lexical meaning. *NeuroImage*, 14, 546–555.
- Janssen, D. P., Roelofs, A., & Levelt, W. J. M. (2002). Inflectional frames in language production. *Language and Cognitive Processes*, 17, 209–236.
- Jeschkiak, 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.
- Kiyosawa, M., Inoue, C., Kawasaki, T., Tokoro, T., Ishii, K., Ohyama, M., Senda, M., & Soma, Y. (1996). Functional neuroanatomy of visual object naming – a PET study. *Graefes Archive for Clinical and Experimental Ophthalmology*, 234, 110–115.
- Kosslyn, S. M., Alpert, N. M., & Thompson, W. L. (1995). Identifying objects at different levels of hierarchy: a positron emission tomography study. *Human Brain Mapping*, 3, 107–132.
- Kosslyn, S. M., Alpert, N. M., Thompson, W. L., Chabris, C. F., Rauch, S. L., & Anderson, A. K. (1994). Identifying objects seen from different viewpoints: a PET investigation. *Brain*, 117, 1055–1071.
- Kuperberg, G. R., McGuire, P. K., Bullmore, E. T., Brammer, M. J., Rabe-Hesketh, S., Wright, I. C., Lythgoe, D. J., Williams, S. C. R., & David, A. S. (2000). Common and distinct neural substrates for pragmatic, semantic, and syntactic processing of spoken sentences: an fMRI study. *Journal of Cognitive Neuroscience*, 12, 321–341.
- Lennox, B. R., Park, S. B. G., Medley, I., Morris, P. G., & Jones, P. B. (2000). The functional anatomy of auditory hallucinations in schizophrenia. *Psychiatry Research: Neuroimaging*, 100, 13–20.
- 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. (1996). Perspective taking and ellipsis in spatial descriptions. In L. Bloom, M. A. Peterson, M. F. Garrett, & L. Nadel (Eds.), *Language and space*. Cambridge, MA: MIT Press.
- Levelt, W. J. M. (1999). Models of word production. *Trends in Cognitive Sciences*, 3, 223–232.
- Levelt, W. J. M. (2001). Relations between speech production and speech perception: some behavioral and neurological observations. In E. Dupoux (Ed.), *Language, brain, and cognitive development: essays in honor of Jacques Mehler*. Cambridge, MA: MIT Press.
- Levelt, W. J. M., Praamstra, P., Meyer, A. S., Helenius, P., & Salmelin, R. (1998). An MEG study of picture naming. *Journal of Cognitive Neuroscience*, 10, 553–567.
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1–38.
- 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.
- Lupker, S. J. (1982). The role of phonetic and orthographic similarity in picture-word interference. *Canadian Journal of Psychology*, 36, 349–367.
- Maddock, R. J., & Buonocore, M. H. (1997). Activation of left posterior cingulate gyrus by the auditory presentation of threat-related words: an fMRI study. *Psychiatry Research: Neuroimaging*, 75, 1–14.
- Maess, B., Friederici, A. D., Damian, M., Meyer, A. S., & Levelt, W. J. M. (2002). Semantic category interference in overt picture naming: an MEG study. *Journal of Cognitive Neuroscience*, 14, 455–462.
- Malow, B. A., Blaxton, T. A., Sato, S., Bookheimer, S. Y., Kufta, C. V., Figliozzi, C. M., & Theodore, W. H. (1996). Cortical stimulation elicits regional distinctions in auditory and visual naming. *Epilepsia*, 37, 245–252.
- Martin, A. (2000). Functional neuroimaging of semantic memory. In R. Cabeza, & A. Kingstone (Eds.), *Handbook of functional neuroimaging of cognition*. Cambridge, MA: MIT Press.
- Martin, A., Haxby, J. V., Lalonde, F. M., Wiggs, C. L., & Ungerleider, L. G. (1995). Discrete cortical regions associated with knowledge of color and knowledge of action. *Science*, 270, 102–105.

- Martin, A., Wiggs, C. L., Ungerleider, L. G., & Haxby, J. V. (1996). Neural correlates of category-specific knowledge. *Nature*, 379, 649–652.
- Martin, A., Wiggs, C. L., & Weisberg, J. (1997). Modulation of human medial temporal lobe activity by form, meaning, and experience. *Hippocampus*, 7, 587–593.
- Mazoyer, B. M., Tzourio, N., Frak, V., Syrota, A., Murayama, N., Levrier, O., Salamon, G., Dehaene, S., Cohen, L., & Mehler, J. (1993). The cortical representation of speech. *Journal of Cognitive Neuroscience*, 5, 467–479.
- McCarthy, G., Blamire, A. M., Rothman, D. L., Gruetter, R., & Shulman, R. G. (1993). Echo-planar magnetic resonance imaging studies of frontal cortex activation during word generation in humans. *Proceedings of the National Academy of Sciences USA*, 90, 4952–4956.
- McGuire, P. K., Silbersweig, D. A., & Frith, C. D. (1996). Functional neuroanatomy of verbal self-monitoring. *Brain*, 119, 101–111.
- McGuire, P. K., Silbersweig, D. A., Murray, R. M., David, A. S., Frackowiak, R. S. J., & Frith, C. D. (1996). Functional anatomy of inner speech and auditory verbal imagery. *Psychological Medicine*, 26, 29–38.
- McGuire, P. K., Silbersweig, D. A., Wright, I., Murray, R. M., David, A. S., Frackowiak, R. S. J., & Frith, C. D. (1995). Abnormal monitoring of inner speech: a physiological basis for auditory hallucinations. *Lancet*, 346, 596–600.
- Mechelli, A., Friston, K. J., & Price, C. J. (2000). The effects of presentation rate during word and pseudoword reading: a comparison of PET and fMRI. *Journal of Cognitive Neuroscience*, 12, 145–156.
- Mellet, E., Tzourio, N., Crivello, F., Joliet, M., Denis, M., & Mazoyer, B. (1996). Functional anatomy of spatial mental imagery generated from verbal instructions. *Journal of Neuroscience*, 16, 6504–6512.
- Menard, M. T., Kosslyn, S. M., Thompson, W. L., Alpert, N. M., & Rauch, S. L. (1996). Encoding words and pictures: a positron emission tomography study. *Neuropsychologia*, 34, 185–194.
- Meyer, A. S., & Schriefers, H. (1991). Phonological facilitation in picture-word interference experiments: effects of stimulus onset asynchrony and types of interfering stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, 17, 1146–1160.
- Meyer, A. S., Roelofs, A., & Levelt, W. J. M. (2003). Word length effects in picture naming: The role of a response criterion. *Journal of Memory and Language*, 48, 131–147.
- Moore, C. J., & Price, C. J. (1999). Three distinct ventral occipitotemporal regions for reading and object naming. *NeuroImage*, 10, 181–192.
- Mummary, C. J., Ashburner, J., Scott, S. K., & Wise, R. J. S. (1999). Functional neuroimaging of speech perception in six normal and two aphasic subjects. *Journal of the Acoustical Society of America*, 106, 449–457.
- Murtha, S., Chertkow, H., Beauregard, M., & Evans, A. (1999). The neural substrate of picture naming. *Journal of Cognitive Neuroscience*, 11, 399–423.
- Ojemann, G. A. (1983). Brain organization for language from the perspective of electrical stimulation mapping. *Behavioral and Brain Sciences*, 2, 189–230.
- Ojemann, G., Ojemann, J., Lettich, E., & Berger, M. (1989). Cortical language localization in left, dominant hemisphere. An electrical mapping investigation in 117 patients. *Journal of Neurosurgery*, 71, 316–326.
- Papathanassiou, D., Etard, O., Mellet, E., Zago, L., Mazoyer, B., & Tzourio-Mazoyer, N. (2000). A common language network for comprehension and production: a contribution to the definition of language epicenters with PET. *NeuroImage*, 11, 347–357.
- Paulesu, E., Goldacre, B., Scifo, P., Cappa, S. F., Gilardi, M. C., Castiglioni, I., Perani, D., & Fazio, F. (1997). Functional heterogeneity of left inferior frontal cortex as revealed by fMRI. *NeuroReport*, 8, 2011–2016.
- Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., & Raichle, M. E. (1989). Positron emission tomographic studies of the processing of single words. *Journal of Cognitive Neuroscience*, 1, 153–170.
- Petersen, S. E., Fox, P. T., Snyder, A. Z., & Raichle, M. E. (1990). Activation of extrastriate and frontal cortical areas by visual words and word-like stimuli. *Science*, 249, 1041–1044.
- Phillips, C., Pellathy, T., Marantz, A., Yellin, E., Wexler, K., Poeppel, D., McGinnis, M., & Roberts, T. (2000). Auditory cortex accesses phonological categories: an MEG mismatch study. *Journal of Cognitive Neuroscience*, 12, 1038–1055.

- Poline, J.-B., Vandenberghe, R., Holmes, A. P., Friston, K. J., & Frackowiak, R. S. J. (1996). Reproducibility of PET activation studies: lessons from a multi-center European experiment. EU concerted action on functional imaging. *NeuroImage*, 4, 34–54.
- Price, C. J., Moore, C. J., & Frackowiak, R. S. J. (1996). The effect of varying stimulus rate and duration on brain activity during reading. *NeuroImage*, 3, 40–52.
- Price, C. J., Moore, C. J., Humphreys, G. W., Frackowiak, R. S. J., & Friston, K. J. (1996). The neural regions sustaining object recognition and naming. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 263, 1501–1507.
- Price, C. J., Wise, R., Ramsay, S., Friston, K., Howard, D., Patterson, K., & Frackowiak, R. (1992). Regional response differences within the human auditory cortex when listening to words. *Neuroscience Letters*, 146, 179–182.
- Price, C. J., Wise, R. J., Warburton, E. A., Moore, C. J., Howard, D., Patterson, K., Frackowiak, R. S., & Friston, K. J. (1996). Hearing and saying. the functional neuro-anatomy of auditory word processing. *Brain*, 119, 919–931.
- Price, C. J., Wise, R. J. S., Watson, J. D. G., Patterson, K., Howard, D., & Frackowiak, R. S. J. (1994). Brain activity during reading. The effects of exposure duration and task. *Brain*, 117, 1255–1269.
- Pugh, K. R., Shaywitz, B. A., Shaywitz, S. E., Constable, R. T., Skudlarski, P., Fulbright, R. K., Bronen, R. A., Shankweiler, D. P., Katz, L., Fletcher, J. M., & Gore, J. C. (1996). Cerebral organization of component processes in reading. *Brain*, 119, 1221–1238.
- Raichle, M. E., Fiez, J. A., Vidian, T. O., Macleod, A.-M. K., Pardi, J. V., Fox, P. T., & Petersen, S. E. (1994). Practice-related changes in human brain functional anatomy during nonmotor learning. *Cerebral Cortex*, 4, 8–26.
- Rees, G., Howseman, A., Josephs, O., Frith, C. D., Friston, K. J., Frackowiak, R. S. J., & Turner, R. (1997). Characterizing the relationship between bold contrast and regional cerebral blood flow measurements by varying the stimulus presentation rate. *NeuroImage*, 6, 270–278.
- Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, 42, 107–142.
- Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition*, 64, 249–284.
- Roelofs, A. (in press). Error biases in spoken word planning and monitoring by aphasic and nonaphasic speakers: comment on Rapp and Goldrick (2000). *Journal of Memory and Language*.
- Rumsey, J. M., Horwitz, B., Donohue, B. C., Nace, K., Maisog, J. M., & Andreason, P. (1997). Phonological and orthographic components of word recognition. A PET-rCBF study. *Brain*, 120, 739–759.
- Salmelin, R., Hari, R., Lounasmaa, O. V., & Sams, M. (1994). Dynamics of brain activation during picture naming. *Nature*, 368, 463–465.
- Salmelin, R., Schnitzler, A., Schmitz, F., & Freund, H.-J. (2000). Single word reading in developmental stutterers and fluent speakers. *Brain*, 123, 1184–1202.
- Salmelin, R., Service, E., Kiesilä, P., Uutela, K., & Salonen, O. (1996). Impaired visual word processing in dyslexia revealed with magnetoencephalography. *Annals of Neurology*, 40, 157–162.
- Schäffler, L., Lüders, H. O., Dinner, D. S., Lesser, R. P., & Chelune, G. J. (1993). Comprehension deficits elicited by electrical stimulation of Broca's area. *Brain*, 116, 695–715.
- Schiller, N. (1998). The effect of visually masked syllable primes on the naming latencies of words and pictures. *Journal of Memory and Language*, 39, 484–507.
- Schmitt, B. M., Münte, T. F., & Kutas, M. (2000). Electrophysiological estimates of the time course of semantic and phonological encoding during implicit picture naming. *Psychophysiology*, 37, 473–484.
- Schmitt, B. M., Schiltz, K., Zaake, W., Kutas, M., & Münte, T. F. (2001). An electrophysiological analysis of the time course of conceptual and syntactic encoding during tacit picture naming. *Journal of Cognitive Neuroscience*, 13, 510–522.
- Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in production: picture-word interference studies. *Journal of Memory and Language*, 29, 86–102.
- Shergill, S. S., Brammer, M. J., Williams, S. C. R., Murray, R. W., & McGuire, P. K. (2000). Mapping auditory hallucinations in schizophrenia using functional magnetic resonance imaging. *Archives of General Psychiatry*, 57, 1033–1038.
- Shtyrov, Y., Kujala, T., Ilmoniemi, R. J., & Näätänen, R. (1999). Noise affects speech-signal processing differently in the cerebral hemispheres. *NeuroReport*, 10, 2189–2192.

- Simos, P. G., Breier, J. I., Fletcher, J. M., Bergman, E., & Papanicolaou, A. C. (2000). Cerebral mechanisms involved in word reading in dyslexic children: a magnetic source imaging approach. *Cerebral Cortex*, 10, 809–816.
- Snyder, A. Z., Abdullaev, Y. G., Posner, M. I., & Raichle, M. E. (1995). Scalp electrical potentials reflect regional cerebral blood flow responses during processing of written words. *Proceedings of the National Academy of Sciences USA*, 92, 1689–1693.
- 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.
- Stemberger, J. P. (1985). An interactive activation model of language production. In A. W. Ellis (Ed.), (1). *Progress in the psychology of language*, London: Erlbaum.
- Stewart, L., Meyer, B. U., Frith, U., & Rothwell, J. (2001). Left posterior BA37 is involved in object recognition: a TMS study. *Neuropsychologia*, 39, 1–6.
- Tagamets, M.-A., Novick, J. M., Chalmers, M. L., & Friedman, R. B. (2000). A parametric approach to orthographic processing in the brain: an fMRI study. *Journal of Cognitive Neuroscience*, 12, 281–297.
- Talairach, J., & Tournoux, P. (1988). Co-planar stereotaxic atlas of the human brain. Stuttgart: Georg Thieme Verlag.
- Tarkiainen, A., Helenius, P., Hansen, P. C., Cornelissen, P. L., & Salmelin, R. (1999). Dynamics of letter string perception in the human occipitotemporal cortex. *Brain*, 122, 2119–2132.
- Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, 381, 520–522.
- Vandenberghe, R., Price, C., Wise, R., Josephs, O., & Frackowiak, R. S. J. (1996). Functional anatomy of a common semantic system for words and pictures. *Nature*, 383, 254–256.
- Van Turennout, M., Ellmore, T., & Martin, A. (2000). Long-lasting cortical plasticity in the object naming system. *Nature Neuroscience*, 3, 1329–1334.
- Van Turennout, M., Hagoort, P., & Brown, C. M. (1997). Electrophysiological evidence on the time course of semantic and phonological processes in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 787–806.
- Van Turennout, M., Hagoort, P., & Brown, C. (1998). Brain activity during speaking: from syntax to phonology in 40 milliseconds. *Science*, 280, 572–574.
- Veltman, D. J., Friston, K. J., Sanders, G., & Price, C. J. (2000). Regionally specific sensitivity differences in fMRI and PET: where do they come from? *NeuroImage*, 11, 575–588.
- Warburton, E., Price, C. J., Swinburn, K., & Wise, R. J. (1999). Mechanisms of recovery from aphasia: evidence from positron emission tomography studies. *Journal of Neurology, Neurosurgery and Psychiatry*, 66, 155–161.
- Warburton, E., Wise, R. J. S., Price, C. J., Weiller, C., Hadar, U., Ramset, S., & Frackowiak, R. S. J. (1996). Noun and verb retrieval by normal subjects. Studies with PET. *Brain*, 119, 159–179.
- Wheeldon, L. R., & Levelt, W. J. M. (1995). Monitoring the time course of phonological encoding. *Journal of Memory and Language*, 34, 311–334.
- Wise, R., Chollet, F., Hadar, U., Friston, K., Hoffner, E., & Frackowiak, R. S. J. (1991). Distribution of cortical neural networks involved in word comprehension and word retrieval. *Brain*, 114, 1803–1817.
- Wise, R. J. S., Greene, J., Büchel, C., & Scott, S. K. (1999). Brain regions involved in articulation. *Lancet*, 353, 1057–1061.
- Wong, D., Miyamoto, R. T., Pisoni, D. B., Sehgal, M., & Hutchins, G. D. (1999). PET imaging of cochlear-implant and normal-hearing subjects listening to speech and nonspeech. *Hearing Research*, 132, 34–42.
- Zelkowitz, B. J., Herbster, A. N., Nebes, R. D., Mintun, M. A., & Becker, J. T. (1998). An examination of regional cerebral blood flow during object naming tasks. *Journal of the International Neuropsychological Society*, 4, 160–166.