W. Hulstijn, H.F.M. Peters and P.H.H.M. Van Lieshout, editors

Chapter 22

PET RESEARCH IN LANGUAGE PRODUCTION

Peter Indefrey

The aim of this paper is to discuss an inherent difficulty of PET (and fMRI) research in language production. On the one hand, language production presupposes some degree of freedom for the subject, on the other hand, interpretability of results presupposes restrictions of this freedom. This difficulty is reflected in the existing PET literature in some neglect of the general principle to design experiments in such a way that the results do not allow for alternative interpretations. It is argued that by narrowing down the scope of experiments a gain in interpretability can be achieved.

INTRODUCTION

If asked whether there are specific areas in the human brain which subserve language production, most people would have the feeling that this question is too vague. It cannot be readily answered because the processing of written or spoken natural language not only involves different *levels*, such as a word level and a sentence level, but also different *kinds of processing*, such as phonological, lexical, syntactic, and semantic processing. It is highly probable that different neural populations subserve these different aspects of language production. Therefore, it does not seem to be a good idea to run a PET or fMRI study simply scanning subjects while they speak. Nonetheless, it has been done. In a PET study by Tamas et al. (1993), subjects were instructed to report how they spent the previous day. As might have been expected, the resulting brain activations comprise a widespread array of left and right areas of all cortical lobes and the cerebellum. One may call this the neural correlate of 'speech-from-memory', as the authors do, but any further interpretation can only be done in a circular way, such as ascribing memory functions to the observed hippocampal activations, because previous work has shown a relation between hippocampal activation and memory. This may be plausible, but the plausibility does not stem from the study at hand, which hasn't added anything to our knowledge. While it is true that to some extent the interpretation of every brain imaging study relies on the synopsis of results from the whole field for mutual validation and the generation of hypotheses, it is also true that this could not work, if not some study by its design had forced a certain interpretation without recurring to previous results. Obviously, the questions about language production that can be addressed successfully in this sense must be more restricted. For this reason, to date, most studies have focussed on language production on the word level, thereby avoiding the additional processing components involved in sentence (or even discourse) level processing.

GENERATING VERBS

Probably the most widely used word level production paradigm in PET and fMRI research is the Verb Generation Paradigm. It was introduced by Petersen et al. (1988), and in the form described here, first used in 1991 by Richard Wise and colleagues from the Hammersmith group in London. In this paradigm, subjects are auditorily presented with a common noun, such as CAKE and asked to produce subvocally as many appropriate verbs, such as 'bake' or 'eat', as they can think of before they hear the next noun. The paradigm today comes in a number of variations, such as presenting nouns visually or asking the subjects to respond aloud, the core of the task, however, has remained unchanged. The task is considered to involve an externally (by choice of the nouns and their presentation frequency) controlled 'semantic generation', that is, a semantically driven lexical search the result of which is the production of the verb. Since also in natural language the expression of a certain concept requires the search for a lexical item, it can be said that the Verb Generation Task must activate brain areas supporting this component of natural language production.

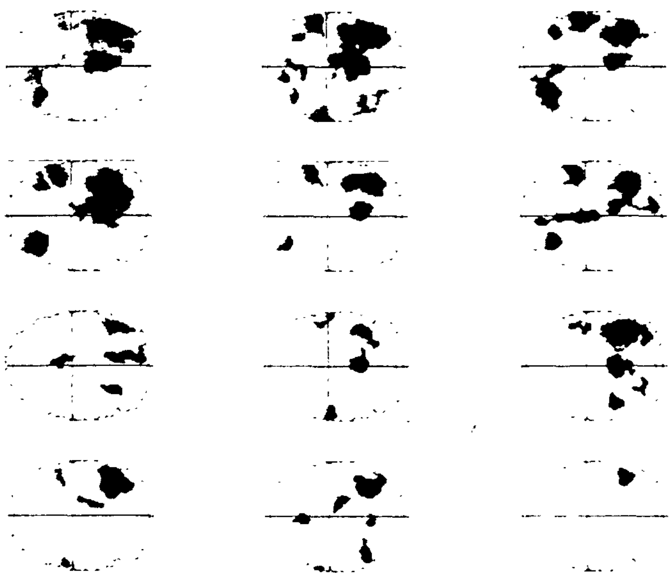

Figure 1. Individual center SPM analyses of a European multicenter PET experiment on verb generation. The subtraction images show transverse views of the significant ($p<0.20$, corrected) regional cerebral blood flow (rCBF) increases during verb generation when compared with rest condition. (adapted from Poline et al., 1996)

In addition to that, it will activate brain areas supporting the subsequent retrieval of the lexically stored phonological word form and probably also secondary motor areas subserving the necessary preparation for spoken output.

The task has proven to be reliable in a recent European multi center study (Poline et al., 1996), for which the stimulus material was translated in seven different languages and subject pretraining, instructions and behavioral assessment were standardized. All data were submitted to a metaanalysis using the statistical parametric mapping (SPM) procedure. Figure 1 shows the high degree of between-center-reliability with respect to the locations of activated regions (centers with small-field-of-view PET cameras in the bottom row).

A further aim of this study was to raise the statistical power by running a pooled analysis over more than 60 subjects in order to reduce the number of false negatives, that is, to assess the number and extent of activated regions that in single center studies fail to become significant. As can be seen in Figure 2, this task does not only activate small specific areas of the brain, but most of the left frontal lobe along with some homologue areas of the right frontal lobe and large parts of the temporal lobes. While it is possible that all these areas indeed subserve the assumed processes, it is also reasonable to ask whether this task did not, after all, comprise additional cognitive components that may have been overlooked. When we look at some of the responses we obtained during the training session from one of the subjects whose data the Nijmegen/Duesseldorf/Juelich center contributed to the European multi center study (see Table 1), it becomes clear that the range of associated verbs cannot be accounted for by assuming a semantically guided search alone. While some of the produced verbs have a 'narrow' semantic relation to the presented noun (BANANA 'peel', TRUMPET 'blow'), many others have a more indirect relation in that they apply to objects in general ('put away').

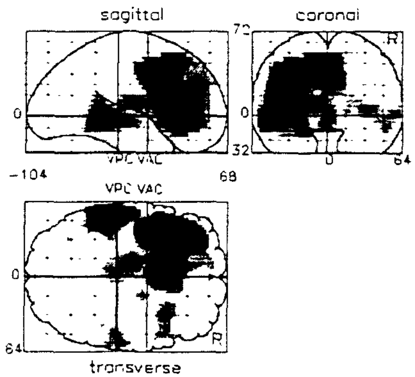

Figure 2. Pooled SPM analysis of the same experiment as in Figure 1 using data of nine centers with a large-field-of-view PET camera. Significant ($p<0.05$, corrected) rCBF increases are displayed in three orthogonal views. (adapted from Poline et al. 1996)

Table 1. Responses to the Verb Generation Task

BANANA	peel, slip on, eat up, plant
TROUSERS	put on, wash, mend, buy, warm
CHAIR	sit, build, nail, sell, work, learn
GLASSES	clean, put on, step on, buy, see
TRUMPET	blow, make music, put away, hear, play
PENCIL	sharpen, break, put away, draw
BUTTON	tear off, close, open
BIRD	fly, eat up, sing
EAR	hear, pinch
DOOR	open, close, kick against

It is particularly problematic that the verbs are not mutually independent, but that they are produced in coherent sequences like 'blow, make music, put away' for TRUMPET or 'build, nail, sell' for CHAIR. These seem to be instances of small scenes that are either recalled from episodic memory or created ad hoc by the subject. In both cases additional processing components would be involved that might contribute to the observed extent of cortical activations. It seems, therefore, as if even a seemingly simple task like verb generation may at closer inspection contain a number of task components that are uncontrolled for. In this situation, it is impossible to attribute specific activation sites to isolated task components on the basis of the task design. It should be added that these problems are shared by a variation of the Verb Generation Paradigm where subjects only produce one verb to every given noun. Of course the episodic nature of verbal associations can then no longer be as easily detected because thematically linked verb sequences do not occur. The problems are also shared by another frequently used task, the Verbal Fluency Task, which is considered to consist in a 'phonologically' driven lexical search. Subjects are asked to produce as many nouns with a given letter as they can think of. While in this task subjects probably are less prone to use some episodic association strategy, they remain nonetheless unrestricted as to how they search for lexical items that meet the condition. Given the letter 'a', for example, some might go for a semantic strategy, producing 'ant, ape, antelope, ...', others might continue to follow the alphabetic route searching for words starting with 'ab...' then 'ac...' and so forth. Thus it remains unclear which additional cognitive processes may be involved in this task. In sum, language production tasks on the word level, although being much more restricted than a "What did You do yesterday?" task, in many cases are still to unrestricted to cope with the flexibility of human speakers on the conceptual level preceding language production.

GENERATING PHONOLOGICAL OUTPUT TO GRAPHEMIC INPUT

It has become clear that the question in functional imaging research is not so much "Can

we *activate* certain cortical regions that are related to a specific cognitive task?" It is rather "Can we *isolate* those cortical regions that subserve the processing component we are interested in?" To give a positive answer to this question, we have to find control conditions that ideally allow for a unique interpretation of the resulting cortical activations. This will often mean that it is better to obtain interpretable results for a small processing component, for which it is possible to design adequate control conditions, than to obtain vague results which, because of a lack of adequate control, can only be interpreted speculatively or in a circular manner.

The processing component for which we tried to isolate the neural correlates in a PET experiment (Indefrey et al., 1995; Hagoort et al., in preparation) is grapheme-to-phoneme conversion. This process is one of two alternative routes to find the appropriate pronunciation for a given written input, the other one being lexical access to a stored phonological representation via a visual input lexicon. The functional distinction between the two pathways is supported by clinical findings (Figure 3 presents the visual-presentation pathways of the model of Patterson and Shewell (1987), which is based on clinical data), as well as by experimental data (see Carr & Pollatsek, 1985 for an overview). Grapheme-to-phoneme conversion relies on the fact that written language, at least in alphabetic and syllabic writing systems, has a non-arbitrary relation to spoken language. This means that it is possible to predict the pronunciation of a word from its spelling. The degree to which this prediction will be correct differs between languages (see Table 2). For languages with a high grapheme-to-phoneme correspondence, like Turkish, knowledge of the conversion rules suffices to pronounce every word correctly even without understanding the language, because there are practically no exceptions to those rules. English, on the other hand, is notorious for its very low grapheme-to-phoneme correspondence, making the pronunciation of an unknown word a risky thing. Just consider the four different pronunciations of the word-final letter string o-u-g-h in the words 'rough, plough, though, cough' or the idiosyncratic pronunciations of 'colonel' and 'gauge'. German, which was used in our PET experiment, lies somewhere between those extremes: correspondence rules will mostly generate the correct pronunciation, but there are exceptional words. Except for languages like Turkish where the writing is almost a phonetic transcription of the spoken language, the term 'grapheme-to-phoneme correspondence rule' is slightly misleading, because these rules are

Table 2. Grapheme-to-Phoneme Correspondence

	Language	Exceptional words
high		
↑	Turkish	none
	German	Agonie [ago'ni:], Begonie [be'go:nj]
		Kabel ['kα:bl], Hotel [ho:'tɛl]
↓	English	rough, plough, though, cough
low		shave, brave, have

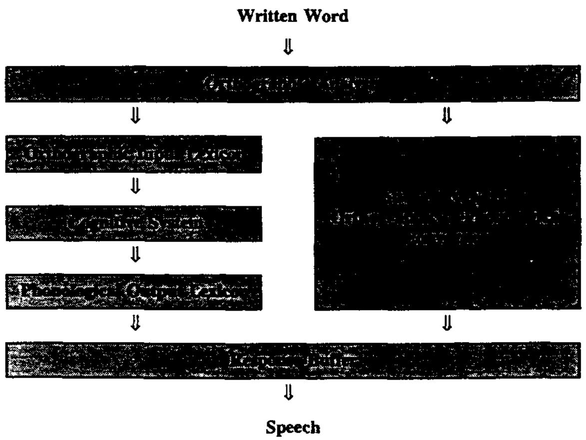

Figure 3. Dual pathway model for the pronunciation of written words. The model is based on clinical data. (adapted from Patterson & Shewell, 1987)

not simple one-to-one mappings of graphemes to phonemes, but rather a set of complex context-sensitive rules. This means that in order to predict the phonemic representation of a certain grapheme one has to inspect the preceding and following grapheme strings. Of course, we are not really dealing with conscious 'prediction' but with tacit linguistic knowledge. The intuitions of speakers of different languages about how to pronounce pseudowords demonstrate the operation of conversion rules (see Table 3). English and German share a rule which states that the quality of a vowel changes depending on whether the following consonant is written as a single or double letter. Exactly in which way the vowel changes, however, is different in the two languages. For Finnish speakers, on the other hand, single or double consonant letters do not indicate the vowel quality but whether the consonant itself has to be spoken short or long.

Grapheme-to-phoneme conversion seems to be an automatic process that is not only active when unknown written words have to be pronounced, but also in silent reading of known words.

Table 3. Pseudoword Pronunciation

	German	English	Finnish
pratte	[-at]	[-æti]	[-at:ɛ]
prate	[- :t]	[-eit]	[-atɛ]

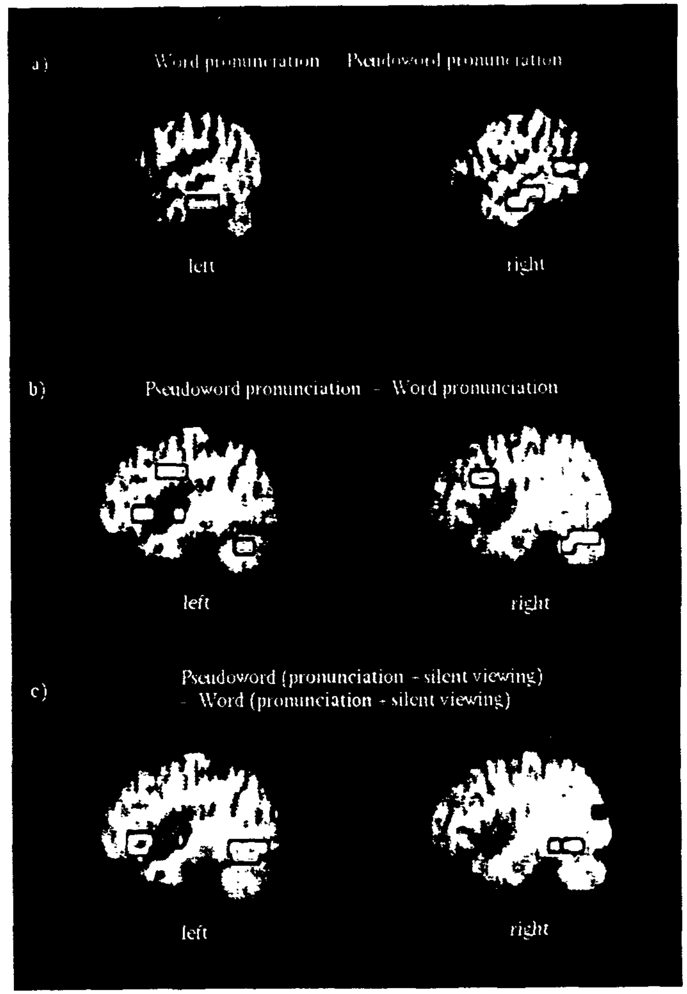

Figure 4. Subtraction images of a PET experiment on the reading of words and pseudowords. Significant ($p<0.05$) rCBF increases in sagittal slices are projected onto corresponding slices of a standardized anatomical MR image. (From Hagoort et al., in preparation)

Thus, subjects need more time to reject visually wrong written sentences that are phonologically correct (e.g. "*He ran threw the streets.*" Baron, 1973; Doctor & Coltheart, 1980).

The crucial point for choosing grapheme-to-phoneme conversion as a processing component that can be successfully isolated in a PET experiment is that it clearly dissociates words and pseudowords. While the production of a phoneme code for words from written input can follow both the lexical and the conversion route, phoneme codes for pseudowords can only be generated through grapheme-to-phoneme conversion. Any additional cortical activation observed for pseudowords must be related to an enhanced recruitment of this process if visual and articulatory complexity are kept constant. In the PET experiment, we directly compared the activations for the conditions (1) Pronunciation of words and (2) Pronunciation of pseudowords. Figure 4 shows the significant activation differences for the subtractions (a) Pronunciation of words - pronunciation of pseudowords and (b) Pronunciation of pseudowords - pronunciation of words. The main finding in (a) is that the pronunciation of words results in a stronger bilateral activation of the middle and inferior temporal gyri than the pronunciation of pseudowords. This confirms previously reported data from other groups (Howard et al., 1992; Price et al., 1994,1996). According to the logic outlined above, however, it does not allow for a unique interpretation, because words have more than one additional property as compared to pseudowords. The observed activations may reflect lexical retrieval of a phoneme code as well as that of a meaning. They may even reflect cognitive processes subsequent to the lexical retrieval. This is different for the main finding in (b), an increase in activation of the left ventral pars triangularis of the inferior frontal gyrus (Broca's area) for pseudoword pronunciation as compared to word pronunciation. Since the only additional property of pseudowords is that their pronunciation relies on grapheme-to-phoneme conversion, the observed activation can be interpreted as being related to this process.

Admittedly, this is an idealized argumentation. The significant motor, premotor and cerebellar activations observed in (b) suggest that pseudowords were also articulatorily more demanding. It is possible to differentiate this motor output-related aspect of pseudowords from the linguistic process by taking into account two further conditions of the experiment, in which subjects read words and pseudowords without overt pronunciation. Although subjects did not speak in these conditions, there was some activation of motor areas relative to a baseline condition in which they silently viewed a fixation cross. This finding suggests that with word-like stimuli the subjects prepared pronunciation to a certain extent without actually executing it, which implies that they also produced a phonemic representation. If we assume that in the silent-reading conditions the motor output processing is more strongly reduced than the cognitive (linguistic) processing, the corresponding cortical activations should be relatively weaker in mean images combining the silent reading and the pronunciation conditions. On the other hand, cortical activations related to grapheme-to-phoneme conversion, which takes place in both conditions, should appear enhanced, due to an improved signal-to-noise ratio. Subtracting the mean image for words from that for pseudowords (c) reveals that the differential activation of Broca's area appears as clearly as before (now together with a bilateral activation of the fusiform gyri which probably reflects some aspect of the visual processing of word-like stimuli which is stronger for pseudowords in both the silent and the pronunciation conditions). Motor, premotor and cerebellar activations, however, no longer reach the significance level. This suggests that they are indeed related to some aspect of articulatory planning that is stronger for pseudowords. There are several possibilities what this aspect might be. A simple explanation might be based on residual segment level differences in the stimulus materials, such as a higher proportion of long as opposed to short vowels in the pseudoword stimuli. It is also possible (and, if true, unavoidable) that the mere fact that a word was novel motivated subjects to articulate it more distinctly. A third, and most interesting explanation would be the following: if motor programs were not only stored for single phonemes but also for syllables (see Levelt, 1989), then the novel words we used in our experiment would not only differ from the known ones in that they required the assembly of a phonemic representation by grapheme-to-phoneme conversion. They would also differ in that they required the assembly of syllable motor programs, because the pseudowords contained a certain proportion of pseudosyllables (all monosyllabic pseudowords, for instance, were by definition at the same time pseudosyllables), which could not possibly have been listed in a memory store. It seems plausible that such an assembly of syllable motor programs might recruit neurons in or near cortical areas known to be involved in motor planning. Interesting as it is, the hypothesis of a memory store for syllable motor programs cannot be confirmed (or rejected) on the basis of the evidence provided by the experiment at hand, simply because the experiment was not planned to confirm it. It nonetheless provides a nice endpoint for this paper, for it is another example of a problem a well designed PET experiment could successfully address, since pseudowords composed out of known syllables and pseudowords composed out of pseudosyllables would differ in just the relevant aspect.

One may object that PET experiments are too invasive, expensive, and time consuming to be devoted to such rather specialized questions. The answer to that is that only PET research on specialized and theoretically well motivated questions provides a chance to slowly create something like a firm ground of knowledge about the neural correlates of cognition.

REFERENCES

Baron, J. (1973) Phonemic stage not necessary for reading. *Quarterly Journal of Experimental Psychology*, *25*, 241-246

Carr, T.H., & Pollatsek A. (1985) Recognizing Printed Words: A Look at Current Models. In Besner, D., Waller, T.G., & MacKinnon, G.E. (Eds.) *Reading Research: Advances in Theory and Practice, Vol.5* (pp. 1-81). New York, Academic Press

Doctor, E., & Coltheart, M. (1980) Children's use of phonological encoding when reading for meaning. *Memory and Cognition*, *8*, 195-209

Hagoort, P., Indefrey, P., Brown, C., Herzog, H., & Seitz, R.J. (in preparation) The Reading of Words and Legal Nonwords: A [^{15}O]-Butanol PET Study.

Howard, D., Patterson, K., Wise, R., Brown, W.D., Friston, K., Weiller, C., & Frackowiak, R. (1992)

The Cortical Localization of the Lexicons. *Brain, 115*, 1769-1782

Indefrey, P., Hagoort, P., Brown, C., Law, I., Herzog, H., Steinmetz, H., & Seitz, R.J. (1995) The Reading of Words and Legal Nonwords: A [^{15}O]-Butanol PET Study. *Human Brain Mapping*, Suppl. *1*, 220.

Levelt, W.J.M. (1989) *Speaking. From Intention to Articulation*. (pp. 326-329) Cambridge (MA), London, MIT Press

Patterson, K., & Shewell, C. (1987) Speak and spell: dissociations and word class effects. In M. Coltheart, G. Sartori, & R. Job (Eds.), *The cognitive neuropsychology of language*. (pp. 273-294) London, Erlbaum

Petersen, S.E., Fox, P.T., Posner, M.I., Minton, M., & Raichle, M.E. (1988) Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature*, *331*, 585-589

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. *NeuroImage*, *4*, 34-54

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

Price, C.J., Wise, R.J.S., & Frackowiak, R.S.J. (1996) Demonstrating the Implicit Processing of Visually Presented Words and Pseudowords. *Cerebral Cortex*, *6*, 62-70.

Tamas, L.B., Shibasaki, T., Horikoshi, S., & Ohye, C. (1993) General activation of cerebral metabolism with speech: a PET study. *International Journal of Psychophysiology*, *14*, 199-208

Wise, R., Chollet, F., Hadar, U., Friston, K., Hoffner, E., & Frackowiak, R. (1991) Distribution of Cortical Neural Networks Involved in Word Comprehension and Word Retrieval. *Brain*, *114*, 1803-1817

SPEECH PRODUCTION: MOTOR CONTROL, BRAIN RESEARCH AND FLUENCY DISORDERS

Editors:

Wouter Hulstijn, Ph.D.
University of Nijmegen
Nijmegen Institute for Cognition Research and Information Technology
Nijmegen, The Netherlands

Herman F.M. Peters, Ph.D.
University Hospital Nijmegen
Department of Voice and Speech Pathology
Nijmegen, The Netherlands

and

Pascal H.H.M. van Lieshout, Ph.D.
University Hospital Nijmegen
Department of Voice and Speech Pathology
Nijmegen, The Netherlands

1997

ELSEVIER

Amsterdam – Lausanne – New York – Oxford – Shannon – Tokyo