The Cambridge Cookie-Theft Corpus: A Corpus of Directed and Spontaneous Speech of Brain-Damaged Patients and Healthy Individuals

Caroline Williams[a], Andrew Thwaites[b], Paula Buttery[c], Jeroen Geertzen[c]
Billi Randall[a], Meredith Shafto[a], Barry Devereux[a], Lorraine Tyler[a]

[a]The Centre for Speech, Language and the Brain, University of Cambridge [b]The MRC Cognition and Brain Sciences Unit
[c]Computation, Cognition and Language Group, RCEAL, University of Cambridge

Abstract

Investigating differences in linguistic usage between individuals who have suffered brain injury (hereafter *patients*) and those who haven't can yield a number of benefits. It provides a better understanding about the precise way in which impairments affect patients' language, improves theories of how the brain processes language, and offers heuristics for diagnosing certain types of brain damage based on patients' speech. One method for investigating usage differences involves the analysis of spontaneous speech. In the work described here we construct a text corpus consisting of transcripts of individuals' speech produced during two tasks: the Boston-cookie-theft picture description task (Goodglass and Kaplan, 1983) and a spontaneous speech task, which elicits a semi-prompted monologue (where the participant answers general non-intrusive questions about their lives and hobbies), and/or free speech (where an initial question is asked and no secondary prompting is required). Interviews with patients from 19*yrs* to 89*yrs* were transcribed, as were interviews with a comparable number of healthy individuals (20*yrs* to 89*yrs*). Structural brain images of 32% of the patients and 18% of the healthy individuals are also available. This unique data source provides a rich resource for future research in many areas of language impairment and has been constructed to facilitate analysis with natural language processing and corpus linguistics techniques.

Introduction

The characteristics of a population's speech can shed light on theoretical models which aim to explain how language is represented and processed in the brain. Typically, these models are based on the phonological, morphological, syntactic and discourse characteristics of language production in young healthy people and their relationship to brain function. Such models provide a baseline against which the language output of patients with brain damage can be evaluated, and can aid in the diagnosis of language impairments (Davis et al., 1998). Moreover, language changes associated both with brain damage and with neural change associated with healthy aging provide strong tests of models of language and brain. (Kemper et al., 2004). However, the development of adequate models and the ability to test them requires input data, in the form of examples of natural speech production, from a wide range of speakers, across the adult lifespan, and from brain-damaged patients with (and without) language deficits. In addition, sufficient data must be collected to allow significance testing of hypotheses based on the transcripts of the speech data. The data should also be richly annotated and easy to manipulate, so that future researchers can readily undertake further analysis of the data. The Cambridge Cookie-Theft Corpus aims to make available this kind of data to the speech and language community.

Background

Research at the Centre for Speech, Language and the Brain [CSLB] aims to explore the language characteristics of brain-damaged patients and possible changes in language as a function of healthy aging. Interviews with patients with specific language disorders (such as syntactic deficits (Moss et al., 1998)) and healthy participants across the adult life-span (Shafto et al., 2007)) have been recorded, providing background data on their naturalistic language use. The raw data from these recordings are highly reusable. The interviews elicit a stream of continuous speech in response to emotionally neutral, open-ended questioning. Questions addressed to patients are designed to make the participant talk about themselves and their interests. In addition, a more constrained set of speech data is obtained by asking participants to describe a picture (in this case The Cookie-theft —see below). Questions take the form of a '*please describe ...*' or '*tell me about...*' query. The collation of speech samples in both naturalistic and constrained contexts can be used to investigate how language production changes due to gradual neural change (i.e. in healthy aging) and punctuate change (i.e. in aphasia).

Data

Computational linguistic techniques can be used to provide statistical analysis of correlations between speech data and neural changes. To facilitate this analysis, our corpus is compiled with the following considerations in mind. The transcription should be, insofar as possible, lossless and error-free. Annotations should be consistent across annotators and across participant interviews. The output should be in a standardized format to enhance its reusability.

These points are addressed in the following ways: The schema for the transcription of the relevant interviews was designed to retain as much information as possible for linguistic analysis, although some information is inevitably lost in the conversion from the audio file to text. Prosody and some phonemic details have not been annotated, though the transcripts have been time-aligned with the audio files, so this information can be added in the future. Inter-annotator variation has been eliminated by using a single annotator for all transcriptions. Consistency across the speech samples is constrained by a rigid annotation guide. This also helps to ensure the consistency of future additions to the corpus. The annotated corpus has been produced in XML format and has been made available with its DTD.

As noted above, the speech samples are derived from a number of interviews of healthy individuals without brain-damage and patients recorded between the mid-1990s and 2009. The interviews were conducted at the CSLB and the MRC Cognition and Brain Sciences Unit [CBU], except for those patients who wished to be interviewed at their homes (sometimes with family members unavoidably present). All healthy individuals were interviewed at the CBU or the CSLB. Insofar as was practical, these recordings were carried out in an isolated environment such as a sound-attenuated interview room.

The data were transcribed using Praat (Boersma and Weenink, 2005) and the output formatted in XML (see Figure fig:xmlexamples for an example). In total, there are 3102 word types and 56288 tokens, covering 9.01 hours of speech. The data were anonymised, (including the suppression of identity-revealing speech content, see page 3), but the speakers in the interviews can be tracked by an ID tag. As well as this ID, each interview transcript has supplementary information attached, including: the patient's diagnosis (i.e., stroke , aphasia, agrammatism, etc.), aetiology (i.e., haemorrhage, infarction, aneurysm, excisions, etc.), area of damage, date of birth, gender and recording date. More information, not yet publically available, includes T1, T2 and DTI scans of a large proportion of the patients and healthy individuals. These structural MR scans were either carried out at the CBU or at the Wolfson Brain Imaging Centre.

Communication content

The interview recordings follow one of two formats. Subjects are either asked questions about themselves or they are asked to describe a picture (the cookie theft). In the first task, the subject is asked questions designed to elicit spontaneous speech, either in the form of a semi-prompted monologue (where the participant answers general non-intrusive questions about their lives and hobbies), and/or free speech (where an initial question is asked and no secondary prompting is required). This task produces a wide range of speech styles, including genuine dialogue, prompted speech, and connected narrative. In the cookie-theft task, the participant is asked to "describe what's going on" in a picture depicting a complex household scene, which includes the notable feature of a child stealing cookies off a high shelf. The cookie theft picture was selected because it is widely used in the study of aphasia(Giles et al., 1996), being included in a popular aphasia diagnostic protocol(Goodglass and Kaplan, 1983). A principal benefit of the cookie theft task is that it allows the variation between speech-styles to be reduced, and also minimizes confounds in analysis due to the controlled nature of the speech content (which is not as controllable in the spontaneous speech tasks). In general, the participant is asked to "*describe what's going on*" in the picture, but the instruction is not always uniform. Although this task is performed without difficulty by all healthy participants, two patients were unable to comply and produced speech that did not always describe the picture. These files can be identified from the fact that even though the task is 'cookie theft', the topic is not.

The length of the patient spontaneous speech samples ranges from 28 seconds to 14 minutes. Most are between 1 -5 minutes long, and are substantially shorter than the recordings from healthy participants (which are typically around 10 minutes duration). There is often less actual speech in a recording due to pausing and false starts. For healthy individuals, each spontaneous speech transcription comprises a two minute extract taken starting from the midpoint of the sound file. The full length of the recording is noted on the transcription, so future researchers know how much is missing. Future work to this corpus will extend this duration. The patient files are similarly truncated, even though on average they are shorter. This ensures maximum coverage of patients despite the time-consuming nature of the transcription task and the difficulties presented by the phonological impairments of some of the patients. The minimum recording duration for patients is 1:51, the maximum is 23:00, and the majority are between 5 and 10 minutes. For healthy individuals, the cookie-theft files are typically 45s long but can range up to 2 minutes, while patient cookie-theft recordings range between 13 seconds and 10 minutes. In these latter cases only 3 minutes were recorded, as these patients were not providing further information on the picture.

Age range	Brain-injured patients		Healthy individuals	
	Cookie-Theft	Spontaneous speech	Cookie-Theft	Spontaneous speech
0-19	0	0	33	3
20-29	5	3	50	12
30-39	7	6	12	4
40-49	14	3	0	3
50-59	18	7	8	5
60-69	22	10	48	9
70-79	18	10	61	8
80-89	3	1	10	1
90-09	0	0	0	0

Table 1: Number of transcriptions per age-range

Participants

The aetiology of patients includes stroke, brain tumours, infarction, haemorrhage, aneurysm, ischaemia, haematoma and medical excisions. The damage is mainly left lateralised focusing on the frontal and temporal cortices; these are thought to be critical to language (Binder et al., 1997). The age range of the patients is 70 years, with the youngest (at the time of recording) 19*yrs* and the oldest 89*yrs*. Table 1 shows the number of interviews transcribed for the corpus. In total there are recordings of 107 cookie thefts from 99 different patients and 129 spontaneous speech recordings from 89 different patients. 78 patients completed both tasks at least once, of whom 41 also have structural brain scans using MRI. The scans provide important additional information to the brief diagnosis provided with each transcription. Patients were selected from a variety of sources: from the neuroscience panel at the CBU (these will normally have detailed medical notes from a clinician), self-referrals, community/self-help groups, or from country-wide memory clinics.

The healthy individuals were volunteers both at the CSBL and the CBU, most of whom were part of a wider panel recruited for other behavioural and neuroimaging studies. There are currently 224 healthy cookie theft recordings and 73 spontaneous speeches, from 232 subjects. T1, T2, DTI scans have been obtained for 32 of the healthy individuals.

Orthographic Transcription

Given our research aims, the transcriptions needed to be easily machine-parsable, and able to treat phenomena systematically, while still providing an adequate representation of speech. The ideal input for many parsers is written text; yet spontaneous-speech transcriptions, if transcribed accurately, are full of information that are not present in written text (fillers, false starts, repetitions, non-standard pronunciation) and implicitly contain much information that is (namely punctuation). The transcription system for this corpus therefore had to negotiate these two (partially-conflicting) goals.

The system was designed and transcribed by the first author (CW). There were no resources for validation (e.g. by evaluating agreement with a second transcriber), therefore the system had to minimise errors from the outset. Given the size of the corpus and the time available, it was decided as far as possible to represent the communication content as a simple transcription without prosodic and phonological information, which is time consuming and can be added in later. Methods for dealing with prosodic and phonological information have been retained (see pages 4 & 5 respectively).

Rational for final transcription schema

The description of the schema (see below) forms an overview of the completed system, and gives an explanation of why certain transcription conventions were used. We consulted the following corpora when designing the transcription system: BASE (developed in accordance with TEI guidelines), Bois et al. (1992), and Bois et al. (1993) were the most influential in the development of the final scheme, with Bois et al. (1993) providing the principles we wanted to use when designing the transcription system.

Due to the nature of the speech (which is in some places heavily fragmented and distorted), our schema is inevitably non-standard. Two corpora that deal with similar speech is the Dutch Corpus of Aphasic Speech (Westerhout and Monachesi, 2006) and PerLA (Paúls, 2004). We often deviate from their approach because they rely on EAGLES and CHILDES whereas we based our guidelines on TEI because it was more XML compatible (i.e. it does not use lots of punctuation characters for annotation purposes) and generally more flexible because it is designed to be modular and adapted.

With the notable exception of prosody (see page 4) as much information as reasonably possible was left in for future linguistic analysis. We based as few transcribing decisions as possible on psychological or linguistic theory. For instance, sentence boundaries and Intonation Points are not recorded. Instead we annotate speaker turns and segments separated by pauses. There are inevitably exceptions to this rule (e.g. suspected truncations —see below), but by and large we followed the same approach as the designers of the PerLA corpus: to treat transcribing and analysis as separate tasks as far as possible, even though the transcription inherently involves judgements about what was said. If enough information is recorded these subjective decisions can be made in a systematic, automated way at a later date. For instance, although immediate research in this area may not specifically look at phonetics or phonology, time-aligning the transcriptions with the audio recordings will allow future researchers to extract this type of information. This time-alignment also gives us rate-of-speech information, tells us how long each sound file lasts, whether and how long pauses are, and tells us where in the sound file a particular feature occurs, but it may also give information about how confident a speaker is at some point, how long the pauses are for word finding difficulties and potentially when sentence boundaries occur.

After the initial transcription is made using Praat, its TextGrid output is converted to a well formed rigid XML schema, which aids later analysis using automated computational techniques.

Linguistic features

Dictionary spellings, abbreviations and contractions were used, in accordance with EAGLES guidelines and the BNC lists where possible. The contractions represent the full spectrum of possible reductions; full-forms are only used if the auxiliary really is completely realised. In addition, filled pauses are kept and lexically transcribed, using a control list amended from Crowdy (1994). Numbers were transcribed in text rather than numerals.

Exact repetitions are marked with `<rep>`. The first use of a word/string is left as is, while subsequent iterations are wrapped in `<rep>` tags with a NO attribute to record which repetition it is (not including the original). `<rep>` can be used for any type of repetition, including phonological, semantic, and syntactic repetition. If the repetition is not an exact repeat then <rep> is not used. Nested repetitions can sometimes be problematic due to the strict XML schema, but these are handled in a systematic way by flagging lexical repetitions at the expense of phrasal ones. Although some sort of psychological judgement is being made here, it is a fairly clear one; in other cases, such as when a participant may be having word finding difficulty, we do not code whether we regard this to be the case psychologically (cf. Westerhout and Monachesi (2006)).

The `<err>` tag is used whenever the speaker appears to identify an error in their speech, for example when they abort a word and try again, for either phonological or semantic reasons, but usually for semantic reasons (phonological errors usually get truncated). Syntactic corrections are handled separately by simply marking truncations. If the speaker does not self-correct then it is not marked with <ERR> tags. A `type` attribute of 'sem' or 'phon' is only added if the nature of the error is unequivocal. The `<err>` tag will have a phonological transcription in <tr> tags inside it (see page 5) if the full word has not been articulated. In these instances the `target` attribute set to the best-guess lexical item (when the word the speaker intended is clear).

Speech fragments for which it is not completely clear what the speaker said are wrapped in `<unclear>` and given a reason, ie. 'distorted phonology'. Like `<err>`, phonological transcription in <tr> tags is often included, ie. when 'distorted phonology' is the reason. If the value is ambiguous (e.g. "*taps*" as plural or "*tap is*") then the `reason` attribute in the `unclear` tag is 'ambiguous'. In cases where what was uttered could not be determined at all, the tag <gap> is used, with the `reason` attribute set to 'inaudible ' or 'unintelligible'.

Segmentation features

As discussed, the speech is time-aligned. It is divided into three tiers comprising of varying resolutions; utterances, segments and sub-segments. The most basic of these tiers are utterances, which we define as: '*a stretch of speech usually preceded and followed by silence or by a change of speaker*' (cf. Traum and Heeman (1997)). Because of this definition, it is important to note that from the point of view of discourse analysis, this is actually closer to the definition of the conversational turn than the utterance, because it is not related to topics or themes (Crookes (see 1990)). Backchannels are always coded as utterances, but they do not inherently trigger end of utterance for the main speaker. However, if the main speaker responds to a backchannel then they trigger an end of utterance, e.g. '*Really?*' '*Yeah, I ...*'. The 'who' attribute is also adopted and is automatically completed in the conversion from TextGrid to XML.

The utterances have been subdivided into segments, where a segment is a stand-alone chunk of text, defined either by pauses, or by the clear rising/falling completion of an intonational phrase. This maintains theory neutrality.

There is a reasonably good correspondence between segments and syntax: segment boundaries are likely to fall at syntactic boundaries in normal speech, although not all syntactic boundaries would have a segment boundary. This would therefore give a subdivision which could be useful to parsers (delimiting syntactically complete strings) while still retaining authenticity in a way that adding punctuation would not.

The lowest span, the sub-segment, does not follow a phonologically acceptable definition and is used merely to keep track of the timing of smaller spans of speech. It covers spans of words that show no change in rhythm, most often between one and four words.

In total, the corpus contains 1331 utterances, 15248 segments, and 18840 sub-segments.

Prosodic features

No prosodic features are recorded, including punctuation. Having said this, we follow Crystal et al. (1989, p59) who notes

```
<file file_id="AB_CBU123_CT_12345678" length="70.3">
 <subject subj_id="AB_CBU123">
  <type>Agrammatic frontal</type>
  <aetiology>Aneursym/ischaemia</aetiology>
  <brain_damage>Left anteromedial temporal pole, LIFG, orbitofrontal, MTG/STG, parietal</brain_damage>
  <dob>1960-01</dob>
 </subject>

 <participants>
  <person role="investigator1" initials="AB" sex="m" />
  <person role="subject" initials="CD" sex="m" />
 </participants>

 <task type="cookie theft" topic="cookie theft" recording_date="2003-03-15">
  <comments></comments>
  <u who="subject" start="0.0" end="52.5">
   <seg start="0.0" end="0.9">
    <subseg start="0.0" end="0.9">erm </subseg>
   </seg>
   <seg start="1.6" end="2.2">
    <subseg start="1.6" end="2.2">mum </subseg>
   </seg>
   <seg start="2.7" end="3.8">
    <subseg start="2.7" end="3.8">washing up </subseg>
   </seg>
   <seg start="5.7" end="6.6">
    <subseg start="5.7" end="6.6">erm </subseg>
   </seg>
   <seg start="6.8" end="8.7">
    <subseg start="6.8" end="8.7">the sink is </subseg>
   </seg>
   <seg start="14.3" end="14.6">
    <subseg start="14.3" end="14.6"><tr>s</tr> </subseg>
   </seg>
   <seg start="19.3" end="20.1">
    <subseg start="19.3" end="20.1"><tr target="flooding">blʌdɪn</tr> </subseg>
   </seg>
  </u>
 </task>
</file>
```

Figure 1: Cookie-theft XML transcription for a brain-damaged patient (abridged biographical details changed)

> ‘it is essential that all speech data is given a transcription of the major prosodic features. [Yet] we do not want too detailed or ‘narrow’ a transcription: there is simply no time to do routine prosodic analysis of a very detailed sort’.

Crystal et al. (1989, p59) thus ‘give systematic recognition only to a subset of prosodic features — to intonation, rhythm and pause.’ As mentioned, rhythm is saved using the time stamped segments (temporal information is given in seconds correct to one decimal place) and pauses are encoded as the space between two segments — later researchers can retrieve pauses from the timings according to their own criteria). Because of their likely co-occurrence with syntactic boundaries, segments broadly capture intonation, albeit not systematically, nor to the detail that Crystal et al. does.

Syntactic truncations are marked with ‘. . .’. This does not imply any kind of pause (unlike the normal written convention for ellipses); contexts in which it is likely to be used are retracings, repetitions, and abandoned or interrupted utterances.

Manner of speech is included, whereby the <shift desc="speech_type"> tag encodes the point at which normal speech has moved to obviously modulated speech such as laughing, reported speech, reading speech, or foreign language speech (with the special attribute lang which states the language). <shift desc="normal" /> signifies the return to normal speech.

Small corpora which implement prosodic transcription do not use punctuation at all and instead use prosodic markers. The BNC uses punctuation to try to ‘approximate to the behaviour of intontational units in terms of pausing and syntax’. COBUILD used trained secretaries to encode normal punctuation. We do not add punctuation because of the theoretical assumptions that this necessarily entails and the lack of systemisation. While intonational markers like accents are not used, we do use macro-level prosody and discourse to break the speech up, by using utterances and intonational phrases to divide below that level.

Phonological information

Phonological information is important, because there are some speakers for whom articulation issues represent a large portion of the impairment. It is also important to retain phonological information so as not to give a misleading impression of what has been said, especially given that one phonological string may have several interpretations.

Phonological transcriptions (<tr target=‘orthographic string’>unicode</tr>) were inserted in the following cases:

a) where the target is unknown but a transcription can be produced (can be surrounded by <gap reason=’unintelligible’> </gap>).

b) where the phonology is non-standard and appears to be a property of the impairment, not part of a dialect (e.g. "cookie gar"). This is often paired with a `<unclear reason='distorted phonology'>` tag, and accompanied by the `target` attribute (see below).

c) where the phonology is non-standard and it is not clear whether this is due to impairment or dialect.

d) for incomplete words or phonemes. These are surrounded by `<trunc>` `</trunc>` and add the target attribute if it is clear what was intended. These truncated words do not trigger repetition tags.

In all cases, the presence of a phonological transcription is a guide to later researchers to pay attention to this section of the text: the addition of analytic information is intentionally avoided.

The `target` attribute tries wherever possible to represent what the speaker intended rather than what they articulated. Specifically, when a speaker with phonological difficulties but apparently no semantic difficulties aims for one word and produces another (e.g. *'off of the stool'* is articulated as *'off of the tool'*), then this will be transcribed as'`<tr target = 'stool'>`tul`</tr>`'. This decision was subjective because the actual word produced could be the result of an error in lexical retrieval rather than an error due to articulation difficulties or faulty phonological representation, but if the speaker otherwise seems to show no problems retrieving semantically appropriate words, this representation is less misleading than putting in a semantically inappropriate word. This also applies to grammatical words: if a patient appears to have mostly intact syntax/morphology but very distorted articulation/phonology, and produces '*he's*' with the vowel of '*his*' then it is transcribed as '*he's*' rather than '*his*', as the latter would imply a grammatical error. Some patients are not tagged with `<tr>` if the articulation difficulties affect every word as it would take too long to transcribe. Instead, this is noted in the XML and such patients should be analysed manually. For the purposes of parsing, only the 'target' attribute of a <tr> can be parsed if needed, so as to avoid clashes where the phonological transcription happens to look like another word.

Miscellaneous

For privacy reasons, identifying names such as personal names and names of home towns/counties were replaced with a `<gap>` *number* `< /gap>`, and set the '`reason`' attribute to '`place`' or '`name`' as appropriate (the '`sex`' attribute is also used if applicable, to facilitate future research of gender agreement for pronouns). The number refers to the referent, thus '*Cathy or Catherine as she was then and I went to the cinema*' would be '*1 or 1 as she was then and I went to the cinema*'.

For non-speech events and external noises, vocal events are recorded as `<vocal>` with description, start, and end times so that rate information can be captured if necessary. They can appear in an utterance or on their own but they are never in an intonation phrase. Inhalations and exhalations are not marked unless extremely salient, because although they are relevant to discourse (Bois et al., 1992), little other prosody is recorded. Non-vocal events (`<incident>`) and gestures (`<kinesic>`) are treated the same way. `<kinesic>` is rare since it is only marked if explicitly recorded on an earlier transcript. The precise label given depends on function rather than realisation: if a grunt sounds like it is functioning as a laugh, then it is labelled as such. Some recordings are made in noisy conditions so only those external noises or actions which are relevant to the discourse are marked.

Fuller details about the transcription system is stored as a technical document in the CSLB.

Future work

There are two current shortcomings in the corpus, both concern the issue of data sparsity. The first is the current gap in ages for healthy individuals with the cookie-theft task between the ages of 25*yrs* and 63*yrs*, for which there are only 33 recordings. The second is a shortfall in the number of instances within each aetiology (for instance, only two patients have semantic dementia, as this, fortunately, is a very rare condition) and damage type. This is due to each of the patients having very different stages and instances of damage. In future, additions to the corpus will focus on these areas.

Acknowledgments

This work is part of the Computational Natural Language Processing and the Neuro-Cognition of Language (CompLex) project, supported by EPSRC (grant EP/F030061/1) and by a Medical Research Council UK grant to LKT (U.1055.04.002.00001.01 and grant G0500842).

References

Binder, J. R., Frost, J. A., Hammeke, T. A., Cox, R. W., Rao, S. M., and Prieto, T. 1997. Human brain language areas identified by functional magnetic resonance imaging. The Journal of Neuroscience:353362.

Boersma, P. and Weenink, D. 2005. Praat: doing phonetics by computer (Version 4.3.01) [Computer program] Retrieved from http://www.praat.org/.

Bois, J. W. D., Cumming, S., Schuetze-Coburn, S., and Paolino, D. 1992. Discourse Transcription, volume 4 of *Santa Barbara Working Papers in Linguistics*.

Bois, J. W. D., Schuetze-Coburn, S., Cumming, S., and Paolino, D. 1993. Outline of discourse transcription, pp. 45–89. *In* J. A. Edwards and M. D. Lampert (eds.), Talking Data: Transcription and Coding in discourse research, chapter 3. Lawrence Erlbaum.

Crookes, G. 1990. The utterance, and other basic units for second language discourse analysis. *Applied Linguistics* 11:193–199.

Crowdy, S. 1994. Spoken corpus transcription. *Literary and Linguistic Computing* 9:25–28.

Crystal, D., Fletcher, P., and Garman, M. 1989. The grammatical analysis of language disability. Cole and Whurr, 2nd edition edition.

Davis, B. L., Jakielski, K. J., and Marquardt, T. P. 1998. Developmental apraxia of speech: Determiners of differential diagnosis. *Clinical Linguistics & Phonetics* 12:p25–45.

Giles, E., Patterson, K., and Hodges, J. R. 1996. Performance on the boston cookie theft picture description task in patients with early dementia of the alzheimers type: missing information. *Aphasiology* 10:395–408.

Goodglass, H. and Kaplan, E. 1983. Boston Diagnostic Aphasia Examination (BDAE). Lea and Febiger. Distributed by Psychological Assessment Resources, Odessa, FL.

Kemper, S., Herman, R., and Lian, C. 2004. Age differences in sentence production. *Journals of Gerontology: Psychological Sciences* 58B:P220–P224.

Moss, H. E., Tyler, L. K., Durrant-Peatfield, M., and Bunn, E. M. 1998. Two eyes of a see-through: Impaired and intact semantic knowledge in a case of selective deficit for living things. *Neurocase: The Neural Basis of Cognition* 4:291– 310.

Paúls, B. G. 2004. La transcripcin del lenguaje afsico, pp. 83–114. *In* B. Gallardo and M. Veyrat (eds.), Estudios de lingstica clnica: Lingstica y patologa. Valncia: Universitat de Valncia - Asociacin Valenciana de Lenguaje, Comunicacin y Cultura.

Shafto, M., Burke, D. M., Stamatakis, E., Tam, P., and Tyler, L. 2007. On the tip-of-the-tongue: Neural correlates of increased word-finding failures in normal aging. *Journal of Cognitive Neuroscience* 19:2060–2070.

Traum, D. R. and Heeman, P. A. 1997. Dialogue Processing in Spoken Language Systems, volume 1236/1997 of *Lecture Notes in Computer Science*, chapter Utterance units in spoken dialogue. Springer Berlin / Heidelberg.

Westerhout, E. and Monachesi, P. 2006. A pilot study for a Corpus of Dutch Aphasic Speech (CoDAS): Focusing on the orthographic transcription. *In* Proceedings of Computational Linguistics in the Netherlands 2005, University of Amsterdam. Amsterdam.