A Re-Evaluation of the Nature of Speech Errors in Normal and Disordered Speakers

Marianne Pouplier[a]   William Hardcastle[b]

[a]Linguistics and English Language, University of Edinburgh, and [b]Speech Science Research Centre, Queen Margaret University College, Edinburgh, UK

Abstract

It is well known that speech errors in normal and aphasic speakers share certain key characteristics. Traditionally, many of these errors are regarded as serial misorderings of abstract phonological segments, which maintain the phonetic well-formedness of the utterance. The current paper brings together the results of several articulatory studies undertaken independently for both subject populations. These show that, in an error, instead of one segment substituting for another, two segments are often produced simultaneously even though only one segment may be heard. Such data pose problems for current models of speech production by suggesting that the commonly assumed dichotomous distinction between phonological and phonetic errors may not be tenable in the current form or may even be altogether redundant.

Copyright © 2005 S. Karger AG, Basel

Introduction

Speech error studies build on the hypothesis that the regularities that can be observed when the system is malfunctioning are the same ones that otherwise govern the production of normal, error-free speech. Methodologically, speech errors have traditionally been collected in the form of written records of slips overheard in normal conversation or in the form of transcription records of (tape-recorded) errors elicited in the laboratory. Results from such research have long been used to argue for the role of abstract phonological representations in speech production. Specifically, errors have been described as arising from a categorical mis-selection that shifts a segment to a wrong position within a prosodic 'frame' [Shattuck-Hufnagel, 1983; Dell, 1986; Levelt, 1989; Berg, 2005], as for instance when the phrase '*a Canadian from Toronto*' is erroneously pronounced as '*a Tanadian . . .*' [Fromkin, 1973]. Unambiguous misorderings of features, such as a nasal in '*cedars of Lebanon => cedars of Lemadon*' [Fromkin, 1971], have only rarely been observed. Meyer [1992] estimates in her overview article that around 60–90% of all errors can be identified as single-segment misorderings, 10–30% constitute sequences of two neighbouring segments (two consonants or a consonant and a flanking vowel), while 'probably less than 5% of all sound

Marianne Pouplier
LEL, University of Edinburgh
Adam Ferguson Building
40 George Square, Edinburgh EH8 9LL (UK)
Tel. +44 131 650 6657, Fax +44 131 650 3962
E-Mail pouplier@ling.ed.ac.uk

errors are feature errors [...]' (p. 184ff.). This has been taken to mean that segments must be direct processing units of word form retrieval, and that syllables and features presumably have only a structural function in speech production [see Roberts, 1975; Guest, 2001, for a different view on feature errors]. While features and syllables have been argued to rarely move as units in errors, they nevertheless seem to play an important, albeit indirect role, in that they influence the likelihood of error occurrence. At the most general level, it can be said that the more two elements have in common, the more likely they are to interact in an error, where the notion of having something in common comprises prosodic position, stress and rhythm, as well as identity of neighbouring segments [e.g. MacKay, 1970; Wilshire, 1999].

Importantly, the auditory impressions of those transcribing the errors suggest that the allophonic properties of a shifted segment appear to be appropriate to its new, erroneous position. This phonetic accommodation has been reported to include aspiration, voicing, vowel tenseness, flapping of /t/ and preceding vowel nasalization [Fromkin, 1971; Garrett, 1975; Stemberger, 1983, 1984]. An example of laryngeal accommodation can be found in Fromkin [1973], where 'steak and potatoes' was heard to be produced as '[speik] and tomatoes' (not *domatoes); that is, the mislocated /p/ and /t/ were (un)aspirated appropriately for their new positions. On the basis of these findings, it has been argued that speech errors arise by shifting the location of abstract segments at a purely phonological planning level, before their phonetic properties are computed. Anomalous execution, on the other hand, has been deemed diagnostic of errors arising in phonetic implementation. For instance, in testing their parallel distributed processing model, Dell et al. [1993] measure the output of their network against 'standards from speech error literature' and assume that 99% of errors are phonotactically regular. The existence of subphonemic errors has always been part of the picture [e.g. Butterworth and Whittaker, 1980; Stemberger, 1983; Berg, 1987], yet these segmentally ill-formed errors have often been set aside under the hypothesis that these errors presumably arise at a different processing stage, outside the reach of phonological regularities, precisely because their non-categorical nature is incompatible with the more commonly observed single-segment shifts.

All the above properties have been taken to corroborate the psychological reality of the segment as the basic unit that is manipulated in accordance with phonological constraints. Most mislocated elements can be described as single segmental misorderings. Features and syllables constrain the occurrence of errors but are not usually error participants themselves. The phonological and phonetic well-formedness of most errors attests that phonological encoding operates on the basis of abstract, symbolic units that are categorically different from the instantiation of these units at a phonetic implementation level.

Errors from speakers with acquired neurogenic disorders have been investigated in a similar vein. In aphasia, a distinction is traditionally made between 'non-fluent' (frequently applied to Broca's aphasia and apraxia of speech, henceforth AOS) and 'fluent' aphasia (such as in Wernicke's and, more controversially, conduction aphasia, the latter described by some [e.g. Geschwind, 1965] as the consequence of an anatomical disconnection between Wernicke's and Broca's areas). Non-fluent aphasia tends to be characterized by speech production deficits accompanied by relatively good auditory processing ability. Speech manifestations may include difficulties in initiating speech, slow, irregular 'groping' articulation, with evidence of sound distortions and what may sound like substitutions, particularly during the production of complex words. AOS is frequently regarded as a separately identifiable disorder [e.g. Wertz et al., 1984],

characterized by difficulties specifically with the phonetic-motor component of speech production, particularly the planning and sequencing of articulatory movements for speech. Fluent aphasics, on the other hand, have impaired auditory processing skills. While their speech is generally said to be fluent and well-articulated, there may however be numerous phonemic paraphasias. These have traditionally been transcribed as serial misorderings (e.g. '*a Tanadian from Toronto*'), additions (e.g. '*fleet*' for '*feet*') or erroneous deletions (e.g. '*gottis*' for '*glottis*') of segments or syllables [Lecours and Lhermitte, 1969; Blumstein, 1973]. In reviewing speech disorder data in this paper, we will be mainly concerned with the so-called 'higher-level' disorders such as aphasia, arising typically from frontal and posterior cortical damage. We will not be covering articulatory impairments such as typically occur in the dysarthrias, involving for example impaired muscular tension, but of course these will certainly need to be accounted for in speech production models [e.g. Weismer et al., 1995]. For a more detailed overview of different types of motor and higher-level speech disorders, the reader is referred to Kent and Tjaden [1997] and Weismer [1997].

It has been pointed out by many researchers that non-transient malfunctions of the speech production process as observed in aphasia share certain key characteristics with transient malfunctions as observed in errors of normal speakers. This has been termed the '*continuity hypothesis*' by Dell et al. [1997b; see also among others Lecours and Lhermitte, 1969; Blumstein, 1973; Buckingham, 1980; Butterworth, 1992; Wilshire, 2002]. For instance, it has been observed that the paraphasias of conduction and Wernicke's aphasia 'all honor the specific segment-ordering conventions of the language' [Garrett, 1984, p. 190], that is, the outcome of an error is phonologically and phonetically well formed. While aphasics' and non-aphasics' speech can be described with a similar error taxonomy, the rate of errors is greater in the former, and some particular characteristics may differ [e.g. Wilshire and McCarthy, 1996]. The relevant point in the present context is, however, that on a more general level, defects arising in aphasia have traditionally tended to be described as defects in phonological processing. Errors of aphasic patients – like those of normal speakers – follow patterns that have been captured within the framework of segmental phonological theories. It should also be pointed out that phonematic deficiencies as observed in fluent aphasia are generally assumed to point to errors in phonological encoding – the process by which word form is retrieved and assembled in the mind – rather than corrupted underlying lexical representations [e.g. Buckingham, 1980; Butterworth, 1992]. It is only under this premise that errors of normal and disordered speakers can be compared. In an aphasic's production of '*Tanadian*' instead of '*Canadian*', it is thus assumed that the underlying lexical representation is correctly '*Canadian*', but the initial /k/ was erroneously substituted by a /t/ during phonological encoding. The basis for this assumption is the observation that the same lexical item will sometimes be produced correctly and sometimes incorrectly.

One area of potential difference between normal and disordered speakers' errors relates to the influence of the context in which the error occurs. Speech errors have been classified as being *contextual* or *non-contextual*, the former denoting an error in which the intruding segment (gesture/feature) comes from neighbouring material (e.g. '*cup of coffee => cuff of coffee*') [Fromkin, 1971], while for the latter, no source for the error can be identified in the immediate neighbourhood of the surface utterance (e.g. '*a three-day thing => a three-way thing*') [Fromkin, 1973]. According to Schwartz et al. [1994], contextual errors account for 70% or more of normal speakers' errors. Talo [1980] reports that 80% of errors are contextual for normal speakers, while less

than 40% of aphasic errors are contextual. We shall return to this as an important point in the final section of the paper.

Many attempts have been made to diagnose different types of neurogenic disorders on the basis of the clustering of different error types [e.g. Blumstein, 1973; Canter et al., 1985; Miller, 1995]. However, the results of these attempts have been largely equivocal which is hardly surprising given the well-attested limitations of relying solely on auditory-based transcriptions. As pointed out by Rosenbek et al. [1984], among others, our traditional view of the nature of disorders such as AOS, originally thought to be predominantly characterized by speech sound substitutions and other phonemic-level errors, may be influenced by the conventions of the phonetic and phonemic transcription techniques we are using to describe the speech data; a point that has also been raised time and again in the context of errors in normal speakers [e.g. Hockett, 1973; Nooteboom, 1969; Cutler, 1981]. While for instance conduction as well as Wernicke's aphasia are indeed assumed to be characterized by phonemic paraphasia, 'subtle phonetic deficits' have been noticed for all types of disorders, so the distinction between the two main types of aphasia (fluent vs. non-fluent) may not be as clear-cut as was originally thought [e.g. Blumstein et al., 1980; Buckingham and Yule, 1987; Vijayan and Gandour, 1995; Dörner and Moriz, 1997].

While the field has in general always been aware of limitations of transcription as a research tool, it has only become evident on the basis of acoustic and articulatory evaluations of anomalous utterances to what extent and in which way the automatic perceptual prefiltering may have provided an incomplete and in some aspects even misleading picture of the nature of errors [e.g. Buckingham and Yule, 1987; Gibbon, 1990; Ingram and Hardcastle, 1990; Mowrey and MacKay, 1990; Scobbie et al., 2000; Pouplier and Goldstein, 2005]. In the last 20 years, several articulatory studies have shown that at least some speech errors may be different in nature than previously thought. While these types of studies have been done with both normal and disordered speakers, and have yielded, as the remainder of this paper will show, comparable results, there has been no discussion so far that has brought together the converging instrumental findings of these subject populations. The new body of data invites us to reassess the validity of the continuity hypothesis and the implications that it may have for models of speech production. The new insights into speech errors will also hopefully have important practical implications, leading to more refined tools in the diagnosis of complex neurogenic disorders.

The goal of the present paper is thus to expand the range of observations that have contributed to the design of speech production models. While there have been acoustic studies of errors – quite numerous for disordered speakers; for normal speakers, the first acoustic study was undertaken by Frisch and Wright [2002; see also Goldrick, in press] – we will confine ourselves in the current paper to a review of articulatory studies, since they render the most direct evidence of what happens in the vocal tract during an error. In the following part, we will refer to the sound affected by an error as the *error target* (in the example above, the /k/ in 'Canadian'), and to the intruding element as *error source* (the /t/ from 'Toronto').

The experimental data available at this point are sparse, especially so in comparison with decades of highly productive transcription-based work. The articulatory investigation of speech errors is hampered by its very asset: the instrumentation employed. The error rate for healthy adult speakers is estimated to be 0.1–0.2% [Garnham et al., 1981] – far too low to obtain more than anecdotal data at best within an experimental

session. Articulatory studies thus by necessity all follow a tightly scripted protocol of stimulus elicitation that triggers errors either by a tongue-twister-based design or by various priming methods. The constraints faced by laboratory studies are non-trivial and sharpen the question about the comparability of naturally occurring and laboratory-elicited errors. Stemberger [1992] particularly cautioned against the use of tongue-twister-based designs in laboratory error elicitation. He states that non-English segments do not occur in naturalistic settings, but are observed in the laboratory. Shattuck-Hufnagel [1983], however, obtained comparable results from corpus and tongue-twister-like error elicitation [see, among others, Sevald and Dell, 1994; Dell et al., 1997a; Wilshire, 1998; Guest, 2001; Goldrick, 2002, for tongue-twister-based designs that study phonological error mechanisms]. Boucher's [1994] observation of spontaneously occurring errors in the laboratory and Pouplier's [2004] results with the SLIP technique are particularly relevant here and will be reported in more detail below. To the extent that the results so far are subject to confirmation and extension in future research, the theoretical considerations have to remain somewhat preliminary, but it is hoped that the current paper will contribute to a fruitful discussion about our current understanding of phonological encoding in speech production.

The finding of transcription-based error research that has proven to be most controversial in the light of instrumental investigations of sound errors is Wells' [1951] 'first law', which concerns the well-formedness of errors. As outlined above, transcription-based research of normal speakers' errors has shown that the vast majority of errors are executed as if intended in their new position; they are phonotactically and articulatorily normal. This in turn has been taken to point to categorically different phonological and phonetic processing stages (and their concomitant error mechanisms) in speech production. In the context of instrumental studies, the question of well-formedness of the outcome of the error has come to be the centre of discussion.

There are, however, a couple of transcription-based studies which specifically examined the nature of so-called ill-formed error outcomes. Laver [1979] in his study on errors during vowel productions identified errors in which the quality of a vowel is between the two targets the subject was instructed to pronounce, and 'of a quite un-English quality' (p. 23). He noticed that of all the vowel pairs presented, only the tense-lax pairs (e.g. '*peep pip*') did not produce errors. He speculates that this is due to the fact that, for these vowel pairs, error source and target are produced with the same muscular system; it will thus be unlikely that conflicting motor commands are triggered. However, vowels differing, for instance, along the front-back dimension, like '*pep poop*', did result in blended pronunciation like 'POEP', with a vowel similar to that in French 'peu', or in a diphthongal pronunciation. Another study by Butterworth and Whittaker [1980], again using a tongue twister elicitation method, found that a substantial number of the errors resulted in 'non-English sequences' (p. 652), such as '*bat gat*' turning into '*gbat gat*', caused by a tendency to add rather than replace. The authors do not, however, draw any further conclusions about the origin of these errors. Stemberger [1983], in his review article, takes care to point out that in corpus data, phonotactic violations in errors are 'not all that uncommon', yet still less frequent than would be expected by chance. The illegal errors he reported consist either of anomalous cluster formation (e.g. '*dlorm*' for '*dorm*') or transposition of the two consonants within a cluster (e.g. '*atk*' for '*act*'). He suggests that errors are more likely to create legal consonant clusters because these, forming a single unit, will necessarily be of higher frequency than illegal clusters and as such more readily accessible.

Overall, very few transcription-based studies have ill-formed errors as their main focus. It has commonly been assumed on the basis of their infrequent occurrences that non-canonical productions are phonetic as opposed to phonological in nature and thus arise at low-level processing stages. We will now turn to the results from articulatory investigations of errors that have been conducted over the last 15 years. This review of existing studies is undertaken with the aim of understanding what the properties of ill-formed errors are and of re-evaluating the continuity hypothesis when drawing the connection to the studies of disordered speech. Further, we will consider how the occurrence of ill-formed errors should be interpreted with respect to the architecture of the speech production process. Generally, well-formedness is taken to be a hallmark of cognitive processes, whereas phonotactic violations are taken to be indicative of motor level processes. Since articulatory studies suggest that ill-formed errors are far more frequent than previously thought, this opens the possibility that the dichotomous distinction between phonological and phonetic errors may not be needed or even not be appropriate. In the context of our understanding of disordered speech, this may have far-reaching implications if instrumentally identified error patterns are to be taken as diagnostic of specific disorders.

Articulatory Studies of Normal Speakers

Articulatory records of errors have been collected by means of electromyography (EMG), electromagnetic midsagittal articulography (EMMA) and X-ray. Overall, these studies have addressed the questions of which units interact in sublexical errors and whether an errorful segment has indeed canonical articulatory properties, as expected on the basis of transcription records. With respect to the latter question, an overarching finding that has emerged from all instrumental studies conducted thus far is that a significantly large number of temporally mislocated segments are atypical in terms of movement amplitude. The first systematic articulatory study of speech errors was published by Mowrey and MacKay [1990] who collected EMG data of the orbicularis oris or the transversus/verticalis complex during the production of tongue twisters such as '*she sells sea shells by the sea shore*'. EMG amplitude variations in anomalous muscle activity ranged from very small to maximal, that is, from just above noise level to an amplitude typical for a normal, error-free production. While Mowrey and MacKay [1990] relied on visual inspection of EMG amplitude, Goldstein et al. [2004] systematically addressed the question of gradience in errors on the basis of EMMA data. Their stimuli consisted of bisyllabic phrases like '*cop top*', repeated in synchrony to a metronome beat. In measuring movement amplitude of the tongue tip and tongue dorsum, the results confirmed the EMG findings of Mowrey and MacKay [1990] that errorful activations occur along a continuum of values, varying from zero to maximal, independent of the type of error occurring.

A seminal paper published by Boucher [1994] examined errors captured on X-ray film for a speaker of French. He discusses, for instance, a case in which by auditory judgment, the segment /ʀ/ seemed to replace /v/: '*trois routes*' (/tʀwaʀut/) was pronounced instead of the intended '*trois voûtes*' (/tʀwavut/). The X-ray data showed, however, that the articulatory movements associated with the labiodental /v/ were still produced, while the /ʀ/ percept resulted from the tongue simultaneously making contact with the uvula, which Boucher interprets as overshooting the position for /u/.

A similar case was an apparent perceptual omission of /ʀ/, despite the presence of the appropriate, yet undershot, articulatory movements. Boucher interprets the errorful over- and undershoot patterns as arising from the movement amplitude specifications of the adjacent vowel. Boucher favours holistic syllable size representations of sound structure, which would allow for coproductions of consonant- and vowel-like features as he observed in the X-ray data. This interpretation goes against the grain of much of speech error research, since usually errors do not occur between adjacent segments, and are thus fundamentally distinct from assimilatory processes conditioned by coarticulation, such as velar fronting. In this context, it may well be that under Boucher's interpretation, these errors should be considered misperceptions rather than mispronunciations. Alternatively, in a more canonical interpretation of the error source, one might want to suggest that the source of the error is not the movement amplitude specification of the adjacent vowel, but the /ʀ/ in the neighbouring word '*trois*'. While this matter cannot be settled here, Boucher's data are truly remarkable since they stem from Rochette's [1973] recordings of French consonants; that is, the data were *not* collected in the context of an error elicitation study. As such, these data are unique and to this date the only published articulatory record of spontaneously occurring errors. That the data consist of two errors from a single speaker again highlights the very essence of the problem articulatory studies face: the error rate in normal speakers is so low that it is next to impossible to systematically study speech errors without any form of error elicitation.

Another intriguing property of errors that has surfaced from the articulatory studies concerns the question of which units interact in errors and, concomitantly, which units should be considered primitive units in speech production. While the predominance of segmental errors has been a longstanding matter of debate [see e.g. Roberts, 1975; Browman and Goldstein, 1990; Levelt et al., 1999; Guest, 2001], again the study by Mowrey and MacKay [1990] was the first to provide articulatory evidence that challenged the predominant segmental view. They suggest that errors are not segmental replacements, but interaction of motor units. Moreover, in errors, 'parts of motor patterns of … two items are produced at the same time' (p. 1311). Yet these errors went unnoticed by the perceiver due to automatic correction mechanisms.

While single motor unit EMG recordings cannot conclusively reveal whether indeed two articulations occur at the same time – muscular counteractivity at an unmonitored site may in fact prevent errorful movement – Goldstein et al. [2004] found that the majority of errors result in a concurrent production of the interacting consonants. For example, in the phrase '*cop top*', in an error a /k/-like tongue dorsum raising would appear during the /t/, *without* the tongue tip raising for the /t/ being omitted; both stops are produced at the same time. That these coproduction errors may occur for articulatory components of segments is suggested by data reported in Goldstein et al. [submitted] on alternations of nasal and non-nasal consonants. In phrases such as '*bad bang*', in some errors the velum and tongue dorsum gesture of the nasal consonant can interact with the coronal gesture for the alveolar stop independently of each other. The authors interpret errors as arising from errorful coordination of gestural units (see below).

Pouplier [2004] extended these findings by using the SLIP technique [Motley and Baars, 1976] to elicit errors while collecting tongue movement data using EMMA. The SLIP technique relies on priming rather than repetition to elicit errors and has thus traditionally been hypothesized to elicit phonological planning errors. Pouplier observed

the same type of coproductions that Goldstein et al. had obtained in their tongue twister experiment, although she observed a significant number of substitution errors as well. Substitutions were defined as a reduced target gesture (e.g. reduction in tongue tip raising during an intended /t/) with concomitant intrusion of another gesture (e.g. tongue dorsum raising, such as would be appropriate for a /k/, during the given /t/). Importantly, both Goldstein et al. and Pouplier point out that substitutions and exchanges could also display movement amplitudes (for tongue tip and tongue dorsum) along a continuum of values. Substitutions as well as coproduction errors could be, but were not necessarily, articulatorily normal in terms of movement amplitude. Both studies treat coproduction errors and substitution errors as the same phenomenon occurring along a continuum of values, similar to assimilation phenomena in fluent speech.

The articulatory investigations of speech errors conducted so far consistently show that a significant percentage of errors cannot be interpreted as symbolic segments replacing each other before their spatiotemporal properties are computed. Anomalous executions in slips of the tongue have appeared in the form of different degrees of movement amplitude (gradience) but also in the form of an intrusion bias, leading to simultaneous productions of the source and the target consonants: errors seem to frequently follow an addition without deletion pattern, in which the source consonant appears, without the intended target consonant being deleted. Yet at the same time, both Goldstein et al. [2004] and Pouplier [2004] also observed errors compatible with the assumption of segmental substitutions, that is, articulatorily well-formed replacements (none of the other articulatory studies address that point). These results have revived the discussion of whether errors of different origin in the speech production process can be differentiated by their articulatory well-formedness (i.e., normal execution). We will return to this question in more detail after the following summary of articulatory studies of aphasic and apraxic speakers.

Articulatory Studies of Disordered Speakers

Most instrumental studies of aphasic and apraxic speech errors have been confined to acoustic analyses examining features such as voice onset time [see reviews in Wambaugh et al., 1996; Auzou et al., 2000; Kent and Kim, 2003]. There have been relatively few instrumental studies looking directly at articulatory processes in aphasia [see summary in Wood and Hardcastle, 2000] and most of these have focused on AOS, presumably because of the on-going controversy surrounding its diagnosis. In the present context, we will review the existing studies again with a main focus on anomalous execution. Contrary to errors in normal speakers, a great deal of attention has been paid to anomalous execution in disordered speech, since this has contributed to identifying the precise nature of the disorder leading to a more informed diagnosis. The key question at hand then is whether we can observe anomalous execution of the same kind that was revealed for errors in normal speakers through the studies reviewed above.

Due to its non-invasive nature, the most popular technique for obtaining articulation records of disordered speakers has been electropalatography (EPG) [Washino et al., 1981; Hardcastle et al., 1985; Sugishita et al., 1987; Hardcastle and Edwards, 1992; Howard and Varley, 1995; Hardcastle and Gibbon, 1997; Southwood et al., 1997; Wood, 1997]. This technique only provides a record of articulations during which the tongue makes contact with the palate; thus, the question of anomalous movement

amplitude is usually not investigated with this technique. One of the earliest instrumental studies was, however, an X-ray microbeam investigation by Itoh and Sasanuma [1984]. They examined the kinematics of articulatory movements of the lower lip, velum and tongue dorsum in an apraxic speaker. Citing evidence of incoordination, both temporal and spatial, between the three different articulators and inconsistency in production, they interpreted this as evidence of faulty programming of speech musculature. They also pointed to the difficulty in labelling the errors they found as either 'distortions' or 'substitutions' when based on perceptual judgments alone.

An early EPG study by Hardcastle et al. [1985] examined tongue-palate contact characteristics of two dysarthric and one apraxic subjects. The results showed that, unlike the dysarthrias, many of the abnormalities revealed by the EPG analysis of the apraxic speaker could be interpreted as errors in the temporal sequencing of phoneme target gestures[1] or as errors in phoneme selection. For example, in the apraxic speaker's attempt at the word '*deer*', a velar stop pattern initiated the word followed by a short period of simultaneous alveolar-velar closure then a relatively long alveolar closure phase. The velar closure pattern which initiated this word was not usually detected by the listener. Also, it did not appear to have been influenced by the context as there were no velars or alveolars preceding or following the word in the corpus. Hardcastle et al. interpreted the intrusive velar gesture as being the result of an error in phoneme selection during the neural sensorimotor planning stage of the utterance. They suggested that this initial selection error of a velar closure is quickly detected by the speaker who attempts to correct it to the required alveolar stop, the temporal transition between the two closures resulting in a short period of overlap, a double articulation. Hardcastle et al. coined the term 'misdirected articulatory gestures' (MAGs) for the intrusion patterns. Hardcastle and Edwards [1992] examined the nature of such MAG errors in more detail in four adult speakers diagnosed as having AOS. While all four speakers showed MAGs, the subjects varied in their timing of the target and the MAG. Hardcastle and Edwards proposed a distinction between the different timing patterns which may have diagnostic significance: an MAG rapidly followed by the target gesture was taken to be indicative of error correction, that is, intact sensory feedback monitoring. A target gesture being rapidly followed by an MAG or a simultaneous production of MAG and target were interpreted as uncorrected errors, and as indicative of a disruption of sensory feedback monitoring.

While these earlier studies suggested that MAGs may be characteristic of AOS, in a larger-scale study, Wood [1997] reports MAGs for eight out of her ten aphasic subjects (two Broca's with AOS and three Broca's without AOS, two conduction and three anomic aphasics). The occurrence of MAGs varied across the subjects. Some, for example one Broca's with AOS, had MAGs in 33.6% of all isolated word productions in the corpus, elicited in a picture-naming task. Alveolar MAGs were found to be almost three times more frequent than velars, yet in a subsequent analysis of Wood's data on sentence production from the same group of aphasics, Dallas [2005] found

---

[1] It should be pointed out that Hardcastle et al. as well as Hardcastle and Edwards use the term '*gesture*' in its general sense, referring to articulator movements in speech production associated with the realization of a particular phoneme. This is different from the more specific use of the term in the framework of articulatory phonology [e.g. Browman and Goldstein, 1989]. Within this framework, a gesture is abstractly defined as a functional grouping of several articulators, that is, it specifies an abstract linguistic task (such as lip closure) rather than a specific articulator movement.

more velar MAGs than alveolars, independent of which type of aphasia had been diagnosed. Wood's data are further remarkable since besides a picture-naming task, she also collected data on spontaneous sentence production. That is, subjects had to use a picture or a target word as a basis for spontaneous sentence production. Wood found a significant number of MAGs in the sentence production task.

The intrusive alveolar gesture here is reminiscent of the type of patterns found by Sugishita et al. [1987] accompanying perceived omissions in their two apraxic speakers. They attributed these MAGs to an inability of the speakers to inhibit alveolar gestures. Wood also undertook an auditory-based analysis using trained transcribers and found that overall, 25% of speech errors categorized as 'pure' substitutions from the auditory analysis involved an additional abnormal lingual-palatal contact when analysed with EPG. Wood's data suggest that MAGs are perhaps far more common in neurogenic disorders than was hitherto thought to be the case and this may explain some of the confusion in the literature surrounding the traditional definitions of AOS. For example Itoh and Sasanuma [1984] noted that most investigators had found substitution errors to be the most prevalent type of articulation in apraxia of speech. However, as they point out in interpretating their X-ray microbeam data, it is distortion errors rather than (phonemic) substitution errors 'that constitute the inevitable consequences of faulty programming of the speech musculature in apraxia' (p. 160).

Overall, studies using techniques such as EMG, EMMA and EPG have found that in speech errors for both subject populations, new gestures are often added without the target gesture being deleted, leading to the simultaneous articulation of two sounds instead of one substituting for another. While full segmental substitutions also occur, the coproduction errors have been observed with significant frequency in disordered as well as normal speakers.

Speech Production Models

In this final part of the paper, we discuss how the data reviewed here might be accommodated by some of the major models of speech production. We will also discuss the question of contextuality of errors, since this factor may be indicative as to where errors arise in the speech production process.

In the light of the data presented here, notably the MAGs and coproduction errors, the question remains how several units of speech production (features, segments or gestures) can come to be encoded at the same time. In Dell's model [Dell, 1986; Dell et al., 1993], phoneme movement errors are due to an erroneous segment having received a higher activation than the target segment. In the selection mechanism, it is not possible for more than one place of articulation to be encoded simultaneously; normally, a single phoneme node with the highest level of activation is tagged and consequently selected at the decision stage. Two phonemes could conceivably share the same level of activation and both be selected, although the model as originally proposed by Dell [1986] does not allow for such a state of affairs. Positive feedback will ensure that the levels of activation keep changing until the point when one phoneme is higher than the others and a decision can be made. While activations in a neural network are of course gradient, and in parallel distributed processing many items have smaller and greater levels of activation, nevertheless the output of the selection process is a sequential string of segments. Conceivably, either activation levels can come to be undifferentiated,

or alternatively the sensitivity of the selection mechanism to differences in activation level could be defunct – although it is difficult to imagine how this would come about in normal speakers. (It could however occur in certain types of pathological conditions affecting speech production.) Alternatively, it could be hypothesized that coproduced consonants are in fact not encoded at the same time but in very rapid succession, which leads to a continuum of temporal overlap values at the surface, one of them being a simultaneous production [see also Hardcastle and Edwards, 1992]. However, since the same type of errors are found in normal and disordered speakers, it is not clear how the Hardcastle and Edwards interpretation of these errors as disrupted sensory feedback monitoring can be extended to normal speakers.

The model by Levelt [1989] and Levelt et al. [1999] does not incorporate feedback between the different stages of encoding – speech production in this feedforward model is strictly serial; what is achieved through direct feedback between levels in Dell's model is ascribed to indirect feedback via a monitoring loop in Levelt's model. Coproduction errors could arise during phonetic encoding, where segmental representations are translated into phonetic gestures. How exactly such a process would work is difficult to judge in the light of the fact that, like Dell's model, Levelt's model does not incorporate any actual articulation. Conceivably, units can interact in the articulatory buffer. According to Levelt [1989] though, '[…] the Phonological Encoder delivers plans for phonological words as smallest units to the Articulator.' How non-contextual errors like the MAGs or an intrusion bias could occur at such a stage is unclear. It is important to note that ill-formed errors, just like segmental substitutions, are not random distortions but can be captured appropriately as interactions of linguistic units smaller than the syllable. While Levelt et al. [1999] explicitly assume that ill-formed errors arise at later, phonetic stages in speech production planning, they also explicitly do not deal with such later stages in their model; ill-formed errors are outside the scope of the current model.

With respect to the question whether anomalous execution should exclusively be assigned to error processes at a phonetic implementation level, non-contextual errors are particularly interesting to consider. Since non-contextual errors are so rare in normal speakers, and are usually not observed under the tightly scripted protocols used in laboratory elicitation when articulatory data are recorded, we can only consider non-contextual errors of disordered speakers here. In Dell's [1986] model, contextual errors occur by virtue of the overlapping fashion in which phonemes are retrieved – more than one phoneme of an utterance are likely to be highly active at the time a given target is selected. Contextual errors arise when a phoneme that is also highly activated is erroneously selected. Errors that are non-contextual at the surface arise in a similar fashion – phonemes from (semantically or phonologically) competing plans can achieve anomalously high activation levels and thus erroneously be selected.

While no articulatory data are available for non-contextual errors in normal speakers, Wood [1997] reported that, in common with many transcription-based studies, around 70% of her coproduction errors were non-contextual. The discrepancy between normal and disordered speakers may of course partly be conditioned by differences in data collection: aphasic patients are usually evaluated on the basis of tasks such as picture naming – an experimental design not used for normal speakers in the context of error studies, since they will perform exceedingly well on that task. It is thus not clear how sharply this discrepancy would emerge under the same set of conditions. Non-contextuality is also difficult to define in a picture-naming task that consists of many repetitions of the same pictures – all items in the list probably come to have a high

baseline activation and thus competition can be expected from other word forms although they are not immediately adjacent. Phonological similarity might then be sufficient in an aphasic speaker to boost activation enough to lead to an error [see Dörner and Moriz, 1997, for a similar point]. The main significance of non-contextual coproduction errors lies in the fact that they are not easily reconcilable with an account in which errors resulting in phonotactically illegal articulations only occur during phonetic encoding or in some kind of articulatory buffer. These errors seem to provide evidence for simultaneous encoding of two phonological units, since non-contextual errors are usually taken to arise from competition in the encoding of alternative utterances [see Wheeldon and Levelt, 1995, for the phonological as opposed to phonetic nature of inner speech, but see e.g. Laver, 1980, for a different approach].

A different model of speech production is the gestural framework [Fowler et al., 1980; Browman and Goldstein, 1986] which allows us to interpret errors within the general context of skilled action. Within this approach, basic units of speech production are assumed to be linguistically significant vocal tract events (termed '*gestures*'), such as the formation or release of constriction. Gestures are modelled on the basis of mass-spring systems [Saltzman and Munhall, 1989] and as such have oscillatory properties. Gestures serve as units of action (constriction formation) and as units of contrast. Within this framework, gestures are hypothesized to be the atomistic units of speech production. Larger, molecular units (comparable to segments or syllables) are constellations of multiple coupled gestures with a particularly high degree of intergestural cohesion [among others, Browman and Goldstein, 1989; Byrd, 1996; Nam and Saltzman, 2003; Goldstein et al., in press]. What sets this model apart from segmental or featural models of speech production – also called translation models [Fowler et al., 1980] – is that gestural lexical representations have inherent spatiotemporal properties. These have also been argued to play a role in grammar [see also Gafos, 2002]. The model does not require an interface stage that translates atemporal symbolic representations into the continuous movement trajectories observed in execution.

In a segmental view, errors during phonological encoding proper will indeed be segmental in nature – segments appear in the wrong serial order *before* their spatial and temporal properties are implemented. Thus, errors resulting in anomalous execution are hypothesized to occur only late in the speech production process [e.g. Levelt et al., 1999; Frisch and Wright, 2002]. Contrary to this, the gestural view predicts that errors will display both gradient as well as categorical properties, independent of where they occur. Due to the spatiotemporal properties of gestural lexical representations, phonological errors can also display anomalous movement amplitude. Further, since gestures are hypothesized to cohere into larger molecular structures (similar to segment size units), the gestural model predicts that the molecular structures as a whole may be involved in errors, while some errors should also affect the individual gestures comprising the larger unit, as was confirmed by the error data on nasal consonants reported in Goldstein et al. [2004].

The only work that discusses errors in aphasia theoretically within a gestural framework is, to our knowledge, the study by Dörner and Moriz [1997]. In their picture-naming (individual words and word pairs) investigation of a patient with mild conduction aphasia, about half of the errors that occurred were identified to be gradient in nature (by acoustic measurements). They interpret this in terms of gestural blending of two word forms being simultaneously encoded. The authors point out that under the assumption of a gestural representation, deficits in phonological encoding will be

manifest in segment size as well as gradient errors. From a gestural point of view, it is thus not necessary to diagnose a secondary low-level disorder in order to account for the non-canonical acoustics they observe in the errors of their patient, since both ill-formed and well-formed properties of errors can be accommodated in the gestural model under a single error mechanism.

How errors in normal speakers can be understood specifically within a gestural framework is elaborated in detail by Goldstein et al. [2004] as well as Pouplier [2004]. The authors propose that coproduction errors are caused by rhythmic synchronization (also called 'entrainment') of coupled oscillators, analogous to the synchronization that is observed in limb oscillation and finger-tapping experiments. Complex coordination tasks will, under certain conditions such as increased rate, show a tendency to converge on a 1:1 frequency, in-phase coordination [Haken et al., 1996; see also Stetson, 1951; Kelso et al., 1986; for a more general picture, see Strogatz and Stewart, 1993]. Goldstein et al. and Pouplier propose that in errors, complex gestural alternations in the course of an utterance come to entrain in a dynamically stable 1:1, in-phase coordination mode (which may or may not correspond to a phonotactically licit constellation) [see Pouplier and Goldstein, 2005]. Errors can then be understood as arising from the interplay of language-specific constraints with extralinguistic dynamic principles known to be characteristic of coordinated movement in general. Gradient properties of errors are part of the critical fluctuations that typically accompany spontaneous transitions between different coordination modes. The entrainment view of errors predicts that coproduction and gradient errors will be observed no matter where in the production process errors arise; there is no dichotomy between 'phonological', well-formed and 'phonetic', ill-formed errors. Errors are instances of dynamic optimization, which may be at the expense of linguistic well-formedness [see also Pouplier, 2003]. How disorders can be accounted for more specifically in such a model and how other characteristics of aphasic speech, such as reduced coarticulation [e.g. Kent and Rosenbek, 1983; Katz, 2000], can be accounted for in detail remains to be addressed. So far, there has also been no attempt to approach a taxonomy of aphasic disorders within a gestural framework. Another important question to be addressed are omission errors. Goldstein et al. [2004] assume these to be part of critical fluctuations, which in a dynamical system characteristically precede the transition to a more stable state. The dynamically more stable state itself is achieved through addition, not deletion (consistent with tapping experiments in which it could be shown that the highest frequency oscillator will dominate) [see Peper et al., 1995]. Within this rhythmic synchronization approach, it will be an intriguing challenge to account for instances of whole-syllable omissions which, according to Cutler [1980], create rhythmic isochrony at the foot level. Also still lacking is a detailed analysis of the temporal characteristics of coproduction errors and the role of phasing in errors, which would help resolve the issue on how two gestures or features can come to be encoded at the same time, or whether they are indeed better thought of as extremely rapid, sequentially encoded segments.

Conclusions

The EPG data discussed above are, in fact, not especially recent – Hardcastle et al. [1985] described MAGs for AOS 20 years ago. It is thus only in the light of the

systematic confirmation of these results by Hardcastle and Edwards [1992] and Wood [1997] for conduction aphasia and Broca's patients with and without AOS and crucially, in conjunction with converging results for normal speakers that these data gain new theoretical importance. To find anomalous execution in errors in AOS may not be surprising, but that similar results were obtained for normal speakers calls for a new theoretical discussion about the implications of anomalous utterances for models of speech production. The limitations of relying on auditory evaluations alone become apparent time and again in articulatory and acoustic studies of errors – this gains particular importance in studies aiming to shed light on the nature of speech disorders, since these results will ultimately impact on diagnosis and therapy. We would expect, for instance, that in patients with clear Wernicke's aphasia, so-called phonemic substitutions will occur along a continuum of magnitudes, and what are auditorily clear segmental substitutions will in a significant number of cases be coproductions of error source and target. Results from articulatory studies are indeed parallel to transcription-based work in that errors of normal and disordered speakers are in many ways comparable. Anomalous executions are not confined to errors in disordered speech. They also occur in normal speakers, and these errors are frequently not segmental replacements (although they may be perceived as such). Yet especially since few articulatory studies have been carried out for normal speakers, it is also clear that at this point we only have a partial picture of what exactly happens in errors and what that may tell us in turn about the mental representation of phonological word form and retrieval mechanisms.

Acknowledgements

References

Auzou, P.; Ozsancak, C.; Morris, R.J.; Jan, M.; Eustache, F.; Hannequin, D.: Voice onset time in aphasia, apraxia of speech and dysarthria: a review. Clin. Linguist. Phonet. *14:* 131–150 (2000).

Berg, T.: A cross-linguistic comparison of slips of the tongue (Indiana University Linguistics Club, Bloomington 1987).

Berg, T.: A structural account of phonological paraphasias. Brain Lang. *94:* 104–129 (2005).

Blumstein, S.: A phonological investigation of aphasic speech (Mouton, The Hague 1973).

Blumstein, S.; Cooper, W.; Goodglass, H.; Statlender, S.; Gottlieb, J.: Production deficits in aphasia. Brain Lang. *9:* 153–170 (1980).

Boucher, V.J.: Alphabet-related biases in psycholinguistic enquiries: considerations for direct theories of speech production and perception. J. Phonet. *22:* 1–18 (1994).

Browman, C.P.; Goldstein, L.: Towards an articulatory phonology. Phonol. Yearb. *3:* 219–252 (1986).

Browman, C.P.; Goldstein, L.: Articulatory gestures as phonological units. Phonology *6:* 201–251 (1989).

Browman, C.P.; Goldstein, L.: Representation and reality: physical systems and phonological structure. J. Phonet. *18:* 411–424 (1990).

Buckingham, H.: On correlating aphasic errors with slips-of-the-tongue. Appl. Psycholinguist. *1:* 199–220 (1980).

Buckingham, H.; Yule, G.: Phonemic false evaluation: theoretical and clinical aspects. Clin. Linguist. Phonet. *1:* 113–125 (1987).

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

Butterworth, B.; Whittaker, S.: Peggy Babcock's relatives; in Stelmach, Requin, Tutorials in motor behavior, pp. 647–656 (North-Holland, Amsterdam 1980).

Byrd, D.: Influences on articulatory timing in consonant sequences. J. Phonet. *24:* 209–244 (1996).

Canter, G.J.; Trost, J.E.; Burns, M.S.: Contrasting speech patterns in apraxia of speech and phonemic paraphasia. Brain Lang. *24:* 204–222 (1985).

Cutler, A.: Syllable omission errors and isochrony; in Dechert, Raupach, Temporal variables in speech. Studies in honour of Frieda Goldman-Eisler (Mouton, The Hague 1980).

Cutler, A.: The reliability of speech error data. Linguistics *19:* 561–582 (1981).

Dallas, J.: EPG study of misdirected articulatory gestures in the connected speech of adults with acquired aphasia; unpublished BSc Honors diss. Edinburgh (2005).

Dell, G.: A spreading-activation theory of retrieval in sentence production. Psychol. Rev. *93*: 283–321 (1986).

Dell, G.; Burger, L.; Svec, W.: Language production and serial order: a functional analysis and a model. Psychol. Rev. *104*: 123–147 (1997a).

Dell, G.; Juliano, C.; Govindjee, A.: Structure and content in language production: a theory of frame constraints in phonological speech errors. Cogn. Sci. *17:* 149–195 (1993).

Dell, G.; Schwartz, M.; Martin, N.; Saffran, E.; Gagnon, D.: Lexical access in aphasic and nonaphasic speakers. Psychol. Rev. *104*: 801–838 (1997b).

Dörner, K.; Moriz, M.: Die lautlich-klangliche Struktur der Äußerungen bei flüssiger Aphasie; unpublished PhD diss. Freiburg (1997).

Fowler, C.; Rubin, P.; Remez, R.E.; Turvey, M.T.: Implications for speech production of a general theory of action; in Butterworth, Language production. Volume 1. Speech and talk, pp. 373–420 (Academic Press, London 1980).

Frisch, S.; Wright, R.: The phonetics of phonological speech errors: an acoustic analysis of slips of the tongue. J. Phonet. *30:* 139–162 (2002).

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

Fromkin, V.A.: Speech errors as linguistic evidence (Mouton, The Hague 1973).

Gafos, A.: A grammar of gestural coordination. Nat. Lang. Linguist. Theory *20:* 269–337 (2002).

Garnham, A.; Shillock, R.C.; Brown, G.D.A.; Mill, A.I.D.; Cutler, A.: Slips of the tongue in the London-Lund corpus of spontaneous conversation. Linguistics. (special issue) *19:* 805–817 (1981).

Garrett, M.F.: The analysis of sentence production; in Bower, The psychology of learning and motivation. Advances in research and theory, pp. 133–177 (Academic Press, New York 1975).

Garrett, M.F.: The organization of processing structure for language production: applications to aphasic speech; in Caplan, Lecours, Smith, Biological perspectives on language, pp. 172–193 (MIT Press, Cambridge 1984).

Geschwind, N.: Disconnection syndrome in animals and man. Brain *88:* 585–644 (1965).

Gibbon, F.: Lingual activity in two speech-disordered children's attempt to produce velar and alveolar stop consonants: evidence from electropalatographic (EPG) data. Br. J. Commun. Disord. *25:* 329–340 (1990).

Goldrick, M.: Patterns in sound, patterns in mind: phonological regularities in speech production (Johns Hopkins, Baltimore 2002).

Goldrick, M.: Cascading activation from phonological planning to articulatory processes: evidence from tongue twisters. Lang. Cogn. Processes (in press).

Goldstein, L.; Byrd, D.; Saltzman, E.: The role of vocal tract gestural action units in understanding the evolution of phonology; in Arbib, From action to language: the mirror neuron system (in press).

Goldstein, L.; Pouplier, M.; Chen, L.; Saltzman, E.; Byrd, D.: Gestural action units slip in speech production errors. Manuscript (2004).

Guest, D.J.: Phonetic features in language production: an experimental examination of phonetic feature errors; PhD diss. Urbana-Campaign, Ill. (2001).

Haken, H.; Peper, C.E.; Beek, P.J.; Daffertshofer, A.: A model for phase transitions. Physica D *90:* 176–196 (1996).

Hardcastle, W.; Morgan-Barry, R.A.; Clark, C.J.: Articulatory and voicing characteristics of adult dysarthric and verbal dyspraxic speakers: An instrumental study. Brit. J. Disord. Commun. *20:* 249–270 (1985).

Hardcastle, W.; Edwards, S.: EPG-based descriptions of aphasic speech errors; in Kent, Intelligibility in speech disorders: theory, measurement, and management, pp. 287–328 (Benjamins, Philadelphia 1992).

Hardcastle, W.; Gibbon, F.: Electropalatography and its clinical applications; in Ball, Code, Instrumental clinical phonetics, pp. 149–193 (Whurr, London 1997).

Hockett, C.F.: Where the tongue slips, there slip I; in Fromkin, Speech errors as linguistic evidence, pp. 93–119. (Mouton, The Hague 1973).

Howard, S.; Varley, R.: Using electropalatography to treat severe acquired apraxia of speech. Eur. J. Disord. Commun. *30:* 246–255 (1995).

Ingram, J.; Hardcastle, W.: Perceptual, acoustic and electropalatographic evaluation of coarticulation effects in apraxic speech. Proc. 3rd Austr. Int. Conf. on Speech Sci. Technol., Canberra 1990, pp. 110–115.

Itoh, M.; Sasanuma, S.: Articulatory movements in apraxia of speech; in Rosenbek, McNeil, Aronson, Apraxia of speech: physiology, acoustics, linguistics, management, pp. 135–166 (College-Hill Press, San Diego 1984).

Katz, W.F.: Anticipatory coarticulation and aphasia: implications for phonetic theories. J. Phonet. *28:* 313–334 (2000).

Kelso, J.A.S.; Saltzman, E.L.; Tuller, B.: The dynamical perspective on speech production: data and theory. J. Phonet. *14:* 29–59 (1986).

Kent, R.D.; Kim, Y.-J.: Towards an acoustic typology of motor speech disorders. Clin. Linguist. Phonet. *17:* 427–445 (2003).

Kent, R.D.; Rosenbek, J.C.: Acoustic patterns of apraxia of speech. J. Speech Hear. Res. *26:* 231–249 (1983).

Kent, R.D.; Tjaden, K.: Brain functions underlying speech; in Hardcastle, Laver, Handbook of phonetic sciences, pp. 220–255 (Blackwell, Oxford 1997).

Laver, J.: Slips of the tongue as neuromuscular evidence for a model of speech production; in Raupach, Temporal variables in speech. Studies in honour of Frieda Goldman-Eisler, pp. 21–26 (Mouton, The Hague 1979).

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

Lecours, A.R.; Lhermitte, F.: Phonemic paraphasias: linguistic structures and tentative hypotheses. Cortex *5:* 193–228 (1969).

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

Levelt, W.; Roelofs, A.; Meyer, A.: A theory of lexical access in speech production. Behav Brain Sci. *22:* 1–75 (1999).

MacKay, D.: Spoonerisms: the structure of errors in the serial order of speech. Neuropsychologia *8:* 323–350 (1970).

Meyer, A.: Investigation of phonological encoding through speech error analyses: achievements, limitations, and alternatives. Cognition *42:* 181–211 (1992).

Miller, N.: Pronunciation errors in acquired speech disorders: the error of our ways. Eur. J. Disord. Commun. *30:* 346–362 (1995).

Motley, M.T.; Baars, B.J.: Laboratory induction of verbal slips: a new method for psycholinguistic research. Commun. Q. *24:* 28–34 (1976).

Mowrey, R.A.; MacKay, I.R.: Phonological primitives: electromyographic speech error evidence. J. acoust. Soc. Am. *88:* 1299–1312 (1990).

Nam, H.; Saltzman, E.: A competitive, coupled oscillator model of syllable structure. Proc. 15th Int. Congr. Phonet. Sci., Barcelona 2003, pp. 2253–2256.

Nooteboom, S.: The tongue slips into patterns. Nomen: Leyden studies in linguistics and phonetics, pp. 114–132 (Mouton, The Hague 1969).

Peper, C.E.; Beek, P.J.; van Wieringen, P.C.W.: Multifrequency coordination in bimanual tapping: asymmetric coupling and signs of supercriticality. J. Exp. Psychol. Hum. Percept. Perform. *21:* 1117–1138 (1995).

Pouplier, M.: The dynamics of error. Proc. 15th Int. Congr. Phonet. Sci., Barcelona 2003, pp. 2245–2248.

Pouplier, M.: SLIPs of the tongue: an articulatory study. Manuscript (2004).

Pouplier, M.; Goldstein, L.: Asymmetries in the perception of speech production errors. J. Phonet. *33:* 47–75 (2005).

Roberts, E.W.: Speech errors as evidence for the reality of phonological units. Lingua *35:* 263–296 (1975).

Rochette, C.E.: Groupe de consonnes en français: étude de l'enchaînement articulatoire à l'aide de la radiociné-matographie et de l'oscillographie (Presses de l'Université Laval, Québec 1973).

Rosenbek, J.C.; McNeil, M.R.; Aronson, A.E.: Apraxia of speech: physiology, acoustics, linguistics, management (College-Hill Press, San Diego 1984).

Saltzman, E.; Munhall, K.G.: A dynamical approach to gestural patterning in speech production. Ecol. Psychol. *1:* 333–382 (1989).

Schwartz, M.F.; Saffran, E.M.; Bloch, D.E.; Dell, G.S.: Disordered speech production in aphasic and normal speakers. Brain Lang. *47:* 52–88 (1994).

Scobbie, J.; Gibbon, F.; Hardcastle, W.; Fletcher, P.: Covert contrast as a stage in the acquisition of phonetics and phonology; in Pierrehumbert, Papers in laboratory phonology V: language acquisition and the lexicon, pp. 194–207 (Cambridge University Press, Cambridge 2000).

Sevald, C.A.; Dell, G.: The sequential cuing effect in speech production. Cognition *53:* 91–127 (1994).

Shattuck-Hufnagel, S.: Sublexical units and suprasegmental structure in speech production planning; in MacNeilage, The production of speech, pp. 109–136 (Springer, New York 1983).

Southwood, M.H.; Dagenais, P.A.; Sutphin, S.M.; Mertz, G.J.: Coarticulation in apraxia of speech: a perceptual, acoustic and electropalatographic study. Clin. Linguist. Phonet. *11:* 179–203 (1997).

Stemberger, J.: Speech errors and theoretical phonology: a review (Indiana University Linguistics Club, Bloomington 1983).

Stemberger, J.: Length as a suprasegmental: evidence from speech errors. Language *60:* 895–913 (1984).

Stemberger, J.: The reliability and replicability of naturalistic speech error data. A comparison with experimentally induced errors; in Baars, Experimental slips and human error: exploring the architecture of volition, pp. 195–215 (Plenum Press, New York 1992).

Stetson, R.H.: Motor phonetics: a study of speech movements in action (North-Holland, Amsterdam 1951).

Strogatz, S.H.; Stewart, I.: Coupled oscillators and biological synchronization. Sci. Am. *269:* 102–109 (1993).

Sugishita, M.; Konno, K.; Kabe, S.; Yunoki, K.; Togashi, O.; Kawamura, M.: Electropalatographic analysis of apraxia of speech in a left hander and in a right hander. Brain *110:* 1393–1417 (1987).

Talo, E.S.: Slips of the tongue in normal and pathological speech; in Fromkin, Errors in linguistic performance: slips of the tongue, ear, pen, and hand, pp. 81–86 (Academic Press, New York 1980).

Vijayan, A.; Gandour, J.: On the notion of a 'subtle phonetic deficit' in fluent/posterior aphasia. Brain Lang. *48:* 106–119 (1995).

Wambaugh, J.L.; Doyle, P.J.; Kalinyak, M.M.; West, J.E.: A critical review of acoustic analyses of aphasic and/or apraxic speech. Clin. Aphasiol. *24:* 35–63 (1996).

Washino, K.; Kasai, Y.; Uchida, Y.; Takeda, K.: Tongue movement during speech in a patient with apraxia of speech: a case study; in Peng, Current issues in neurolinguistics: a Japanese contribution. Proc. 2nd ICU Conf. on Neurolinguist., Tokyo 1981, pp. 125–159.

Weismer, G.: Motor speech disorders; in Hardcastle, Laver, Handbook of phonetic sciences, pp. 191–219 (Blackwell, Oxford 1997).

Weismer, G.; Tjaden, K.; Kent, R.D.: Can articulatory behavior in motor speech disorders be accounted for by theories of normal speech production? J. Phonet. *23:* 149–164 (1995).

Wells, R.: Predicting slips of the tongue. Yale Sci. Mag. *26:* 9–30 (1951).

Wertz, R.T.; LaPointe, L.L.; Rosenbek, J.C.: Apraxia of speech in adults: the disorder and its management (Gruner and Stratton, Orlando 1984).

Wheeldon, L.; Levelt, W.: Monitoring the time course of phonological encoding. J. Mem. Lang. *34:* 311–334 (1995).

Wilshire, C.E.: Serial order in phonological encoding: an exploration of the 'word onset effect' using laboratory-induced errors. Cognition *68:* 143–166 (1998).

Wilshire, C.E.: The 'tongue twister' paradigm as a technique for studying phonological encoding. Lang. Speech *42:* 57–82 (1999).

Wilshire, C.E.: Where do aphasic phonological errors come from? Evidence from phoneme movement errors in picture naming. Aphasiology *16:* 169–197 (2002).

Wilshire, C.E.; McCarthy, R.A.: Experimental investigations of an impairment in phonological encoding. Cogn. Neuropsychol. *13:* 1059–1098 (1996).

Wood, S.: Electropalatographic study of speech sound errors in adults with acquired aphasia; unpublished PhD diss. Edinburgh (1997).

Wood, S.; Hardcastle, W.: Instrumentation in the assessment and therapy of motor speech disorders: a survey of techniques and case studies with EPG; in Papathanasiou, Acquired neurogenic communication disorders: a clinical perspective, pp. 203–248 (Whurr, London 2000).