Introduction

Carson T. Schütze
University of California, Los Angeles

Victor S. Ferreira
University of California, San Diego

1. Motivation

1.1 Goals of the workshop

The purpose of the workshop commemorated in this volume was to assemble leading researchers in the field of speech errors (a.k.a. slips of the tongue) to exchange ideas in conjunction with a wider group of linguists of varied interests and backgrounds gathered for the Linguistic Society of America Institute (co-hosted by MIT and Harvard University in the summer of 2005). The goals of doing this were

> 1) to provide a concentrated opportunity (which there had not been in quite some time—see section 1.2 for those interested in speech error issues to share knowledge of new developments and jointly shape future directions for research;
>
> 2) to catalyze the synthesis of tried-and-true as well as newer speech error techniques with emerging methodologies in the cognitive sciences, including tools from neuroscience and computational science;
>
> 3) to promote interdisciplinary interaction by exposing recent exciting developments from speech error research to the wider linguistics community, which is witnessing a broadening of methodological approaches, and for whom a new wave of speech error work is considerably more relevant than the classic studies that most linguists may be aware of;
>
> 4) to solicit input from various linguistic domains that can inform this new line of speech error research, in order to take greater advantage of the understanding of human language gained since the versions of linguistic theory that many psychologists were exposed to in their training.

The rest of this introduction situates the workshop in its historical context and briefly outlines how the contributions to this volume can be seen as fitting together.

MIT Working Papers in Linguistics 53, 1–7
The State of the Art in Speech Error Research: Proceedings of the LSA Institute Workshop
Carson T. Schütze and Victor S. Ferreira (Eds.)

1.2 Why this meeting now?

The only previous meeting we are aware of that was devoted specifically to this topic was a Working Group on Speech Errors held at the XIIth International Congress of Linguists in Vienna, Austria in 1977 and organized by Victoria Fromkin. Twenty-seven pre-circulated papers were discussed over one-and-a-half days (Fromkin 1978). A collection of most of the papers presented there was published (Fromkin 1980) and continues to be widely cited. Even that meeting included discussion of perceptual errors (so-called slips of the ear), however, so our workshop was in fact the first of its kind in terms of focus. Slip research remains an important area of current psycholinguistic inquiry. This is witnessed by the fact that a large (but surely still rather incomplete) bibliographical search identifies nearly 100 works published since the year 2000 primarily concerning speech errors. This observation, in conjunction with the fact that a meeting focused on speech error work had not been held in nearly 30 years, pointed to a pressing need to provide a forum for those interested in this topic to exchange insights, opinions and advice.

Moreover, we see a re-birth underway of the contribution of speech errors to language (processing) research, not simply in the way they were used in the 1970s, but as part of more sophisticated investigations that take full advantage of the technologies and insights that cognitive science has accrued in the intervening decades. A new generation of psycholinguistic researchers have recently begun independent careers, a generation exposed to very different conceptual frameworks and paradigmatic concerns from those of 30–35 years ago. The workshop, and this volume, have provided opportunities to meld the fresh ideas of this new generation with the knowledge and experience of the pioneers in the field, in order to promulgate and advance the state of the art in speech error research.

2. A non-impartial partial overview of speech error research

Since at least the late 1960s, speech errors have served as the primary source of evidence for how the real-time language production system of the human mind works. This central role was the result of a methodological conundrum: whereas in studying comprehension we are dealing with a mapping from a manipulable external signal (e.g. speech) to a mental representation, in production the process starts with a mental representation, which is rather difficult to manipulate without also determining what the output of the mapping, namely the verbal expression of that mental concept/intention, will be. Thus, it seemed that our best hope for studying the production system might be to study how it breaks down in everyday language situations and try to deduce its architecture from patterns of errors. This endeavor, based on naturalistically-collected corpora of spontaneous slips of the tongue, was remarkably successful, but also always limited by certain methodological concerns, especially the potential for sampling bias in the errors that people would bother to write down for inclusion in their corpora (Bock 1996). The field got a big boost in the mid-1970s with the development of a technique for eliciting (a particular kind of) slips of the tongue in the laboratory (see Baars 1980). Relative rates of slips in this paradigm turned out to make theoretical sense and to often correlate with patterns in spontaneous slips (Stem-

berger 1992), although concerns about the ecological validity of the technique remain (again, see Bock 1996). Through the 1970s and into the 1980s, dozens of generalizations about speech errors were uncovered by these two methods, and considerably detailed models of the production system were constructed to account for them (for review see Humphreys 2002).

Things began to change in the later 1980s and into the 1990s, largely due to the work of Pim Levelt and his colleagues at the Max Planck Institute in Nijmegen. Levelt felt that constructing models of production entirely on the basis of error data was risky—how certain can we be that the way things go *wrong* accurately reflects what happens when they go *right*?—and essentially introduced a whole new set of experimental methods, based on reaction times in various (typically error-free) tasks (for discussion, see Levelt et al. 1999). From this new kind of experimental evidence, Levelt built up a comprehensive model of all stages of language production, and reduced the role of speech error data in the field. However, Levelt's model does not provide comprehensive coverage of the data in the speech error record, and of course, the key insights about language production that came from two decades of speech error research have not disappeared. This, in our view, indicates that the methodological pendulum in the field swung too far, and it is now time for it to swing (part-way) back.

Furthermore, the concerns that led to the emergence of speech error research in the first place are still with us. The picture-naming tasks, paired associate tasks (where speakers are instructed to say *cat* each time they see the prompt *dog*), and translation tasks that form the core of the chronometric tradition advocated by Levelt and colleagues have important limitations. For example, picture-naming tasks can only elicit descriptions of concrete entities; paired-associate tasks are subject to strategic effects; translation tasks require controlling for two languages' worth of potential confounds. The necessity for converging evidence in any domain of investigation is as pressing today as it was in the 1960s; the objective of the workshop and of this volume is to inform the linguistic and psycholinguistic communities of emerging developments in the area of speech error research that, in conjunction with other newly developing techniques, will enable this work to go beyond its previous boundaries and continue to contribute importantly to our understanding of language production and grammar.

3. Contributions to this volume

The papers emanating from presentations at the workshop have been loosely grouped into three sections in this volume. The first eight constitute invited talks and commentaries thereon. The following eleven arise from other talks and posters submitted on specific research programs. The final three contributions look at the broader picture.

The first section of the volume proceeds from contributions exploring "higher-level" language issues to those exploring "lower-level" language issues—beginning somewhat deep inside the head, if you will, and ending almost at the articulators. The section begins with a contribution by Roland Pfau that strongly exemplifies the approach we hoped to foster with the workshop. Pfau takes a commonly investigated phenomenon in the speech error literature, ac-

commodation (e.g., as the result of an exchange, *You're too good for that* becomes *That's too good for you* rather than *That're too good for you*), and analyzes it in terms of an innovative formal linguistic framework called Distributed Morphology, arguing that the architecture of this model largely does away with the need for processing operations designed solely to implement accommodation. Adam Albright's commentary then extends a specific aspect of Pfau's proposal to gain further theoretical insights within the Distributed Morphology framework. The next contribution, from Thomas Berg, uses speech error data to argue for cross-linguistic differences in the internal hierarchical structure of the syllable (or lack thereof). Continuing the metrical theme, Gary Dell and Jill Warker present a series of experiments that explore constraints on speech errors based on syllable position, identifying effects of experience within an experimental session upon the patterns of errors that are observed. They use these experiments to motivate a computational model of such learning effects. Matthew Goldrick's commentary generalizes Dell and Warker's analyses and models, resulting in an insightful unification of their psycholinguistic evidence with formal linguistic theory (specifically, certain forms of Optimality Theory).

The first section of the volume then concludes with a trio of contributions exploring a fundamental issue in the language sciences, namely, whether human linguistic representations are fundamentally categorical or fundamentally continuous and gradient in nature. Marianne Pouplier presents results of experiments using articulatory measurements to conclude that many speech errors involve blending of alternative speech gestures, resulting in vocal tract configurations that do not occur in error-free speech and thus revealing a noncategorical nature to the participating linguistic representation. Joe Stemberger presents a contrasting view, using acoustic analyses of speech errors to argue that, in a critical way, error patterns display a categorical nature and do not lie outside the range of error-free productions. Furthermore, he points out that it is no longer obvious what it means for something to be an error at all. Addressing both of these positions, Stefan Frisch concludes the section by analyzing the differences between the previous articulatory versus acoustic investigations, contextualizing some of their opposing conclusions in terms of different levels of analyses (linguistic vs. cognitive, competence vs. performance) to better understand how they truly do (or do not) conflict.

The second part of the volume begins with a series of papers that investigate various sentence- or phrase-level phenomena that arise in speech-error research. The first two explore the oft-investigated area of agreement and coreference. The first contribution, from Julie Franck, Uli Frauenfelder, and Luigi Rizzi again represents well the synthetic goal of the workshop. Franck et al. report a wide-ranging and systematically varied data set investigating subject–verb agreement in different structures. The resulting overall pattern across the experiments is surprisingly orderly, and supports a number of specific claims within current theories of syntactic structure. Next, Bob Slevc, Liane Wardlow Lane, and Vic Ferreira present a pair of more traditional psycholinguistic studies that aim to investigate whether the interference observed in establishing the relationship between a genitive pronoun and its antecedent is best described as arising from semantic or meaning-level ("world") knowledge versus linguistic/lexical-level ("word") knowledge. The three papers that follow explore other structural aspects of phrasal knowledge and phrase construction. Noriko Iwasaki

explores case-particle errors in Japanese, gaining insight on the possible default status of the nominative (*-ga*) particle (and drawing some methodological lessons along the way). Wardlow Lane, Slevc, and Ferreira report a meta-analysis of two series of experiments, aiming to identify the linguistic and methodological factors that encourage exchange errors and syntactic accommodations. Liz Coppock presents an innovative computational analysis of a broad corpus (of somewhat unusual origin) to determine whether blend errors (e.g., *not letting the pressure on*, which blends *not letting the pressure up* and *keeping the pressure on*) are better analyzed as being constrained by syntactic alignment between the blending alternatives, vs. by a prosodic alignment anchored by the location of stress.

The second group of papers then continues with three contributions exploring issues broadly related to lexical access. Belen Lopez Cutrin and Gabriella Vigliocco report a study of the perhaps most maddening form of speech error, the infamous tip-of-the-tongue state (when speakers know they know a word but cannot produce it). In addition to confirming the previous important observation that speakers in tip-of-the-tongue states can nonetheless report the grammatical gender of the elusive word (even in the absence of phonological information about that word), they also describe an intriguing "tip-of-the-mind" state that involves retrieval of knowledge halted at a conceptual level of representation. The next contribution, from Julio Santiago, Elvira Pérez, Alfonso Palma, and Joe Stemberger, describes a surprising effect with contextual errors (at the word level, and also at syllable- and phoneme-levels). Specifically, they reveal that the representations that participate in such errors tend to be lower frequency than chance. Most surprisingly, source representations tend to replace target representations of higher frequency (what they call the David effect), implicating imperfect or sloppy processing of the source as a key cause of errors. Then, Lise Menn and Michael Gottfried use a creative and valuable technique for eliciting lexical selection errors from aphasic patients and unaffected controls: speakers are presented with scenes in odd configurations (e.g., a chair facing backwards toward a table) and asked to describe them. With this technique, a wide array of forces that influence speakers' selection errors can be identified.

Two studies then use cross-linguistic comparisons to gain insights into speech error phenomena. The first, by Helen Leuninger, Annette Hohenberger, and Eva Waleschkowsi, employs two languages that are about as different as they can be: Spoken German and Deutsche Gebärdensprache (DGS, or German Sign Language). By exploring the nature of the differences between two languages that are articulated and perceived by almost completely different peripheral mechanisms, truly core aspects of human language are better revealed. Most interestingly, Leuninger et al. use the basic differences between Spoken German and German Sign Language to better understand what would otherwise be a striking difference between patterns of speech errors observed in the two languages. They argue for a principled distinction between error patterns that reflect properties of the processor (which are expected to be universal) versus those that reflect properties of the language being processed (which are expected to show variation), and demonstrate how this dichotomy plays out in the two languages they compare. The second contribution, from Jeri Jaeger, compares well-known findings from Germanic languages with lesser-known data from non-Germanic ones to explore what was thought to be a general (and perhaps

universal) trend with speech errors, namely, that anticipatory errors are more common than perseveratory errors. The non-universality of this asymmetry leads Jaeger to suggest that other differences among the languages, having to do with prosodic and information-structure characteristics, are responsible for the relative rates of the two kinds of errors.

The final contribution to this second section provides a useful tool for making quantitative speech error claims. In this contribution, Joe Stemberger presents logical analyses and mathematical derivations to show that rates at which different kinds of errors are observed can be directly compared despite some challenges to making such comparisons (having to do with differences in base rates or opportunities for errors), by comparing the observed rates of errors scaled to chance estimates.

The third and final group of papers take a wider view on methodological, conceptual, and empirical issues. Stefanie Shattuck-Hufnagel's paper stems from her invited talk that introduced the panel discussion on the status of traditional speech error corpora as research tools in the future. The editors' report on that panel session follows. For the final talk of the workshop and the final paper of this volume, we invited Merrill Garrett to evaluate the state of the art in speech error research, as reflected in the presentations at the workshop, in the context of where the field has come from and where it may be heading.

References

Baars, Bernard J. (1980). On eliciting predictable speech errors in the laboratory. In Fromkin 1980, 307–318.

Bock, Kathryn (1996). Language production: Methods and methodologies. *Psychonomic Bulletin & Review* 3, 395–421.

Fromkin, Victoria A. (1978). Speech errors: Slips of the tongue, ear, pen, and hand. (Working group 2.) In *Proceedings of the Twelfth International Congress of Linguists*, ed. Wolfgang U. Dressler and Wolfgang Meid, 120–122. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.

Fromkin, Victoria A. (ed.) (1980). *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand*. New York: Academic Press.

Humphreys, Karin R. (2002). Lexical bias in speech errors. Doctoral dissertation, University of Illinois, Urbana-Champaign.

Levelt, Willem J. M., Ardi Roelofs and Antje S. Meyer (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences* 22, 1–38.

Stemberger, Joseph Paul (1992). The reliability and replicability of naturalistic speech error data: A comparison with experimentally induced errors. In *Experimental slips and human error: Exploring the architecture of volition*, ed. Bernard J. Baars, 195–215. New York: Plenum Press.

(Schütze)
Linguistics Department
UCLA
3125 Campbell Hall
Box 951543
Los Angeles, CA 90095-1543
USA

cschutze@ucla.edu
http://www.linguistics.ucla.edu/people/cschutze

(Ferreira)
Psychology Department
UCSD
9500 Gilman Drive
La Jolla, CA 92093-0109
USA

ferreira@psy.ucsd.edu
http://lpl.ucsd.edu/