PREFACE

It has been a long-standing tradition to publish an annual index of *Language* in the final issue of each volume. In addition, the Society has published two collated indexes of *Language*, one covering the first fifty years of the journal (1925-1974) and the other covering the next ten years (1975-1984). Each of these indexes was published separately, serving as a reference for articles and authors, along with book notices and book reviews and their corresponding authors.

It has now been nearly twenty years since a cumulative index of this type was published, though annual indexes have always appeared in the December issue of the journal. The next independently published index to *Language* might easily have been one incorporating the years 1985-2000, capping off the century. Recent advances in technology, though, made it possible to consider preparing a single collated index of the first seventy-six years of *Language*, beginning with the first issue published in 1925 and continuing on to the last year of the millennium, the year 2000. While past indexes were published as hard copy alone, technology has also allowed us to publish this cumulative index in two formats, each with its own advantages: in handy book form and as a searchable electronic database. In this second format, the *Language* index will become a living document, updated each year as the annual index is prepared and made available to members on the Linguistic Society of America’s web page.

The making of this cumulative index was both time-consuming and complicated. Our basic goal of providing a complete machine-readable and searchable alphabetical index of all twentieth-century volumes of *Language* was deceptively simple to achieve in principle, since we had indexes covering the entire century: we needed only to gather together all these indexes of the past and to collate the entries of theses indexes into a single alphabetical list. This apparently simple task proved to be difficult, however, for two major reasons. First was the sheer number of entries involved in an index this size: approximately 25,000 covering the first seventy-six years of the journal. Second and more formidable was creating the electronically manipulable documents that could then be collated into a single index. Creating these electronic documents required considerable effort, since the indexes for all but the last seven years were available only as hard copy and had to be converted into electronic format in four steps: scanning the paper documents in order to convert them into electronic images; converting the scanned images into text documents by means of optical character recognition (OCR); editing these documents electronically into a single unified text; and finally (and most formidable), going through the entire final document manually several times, searching for characters that OCR had either failed to recognize or, more insidiously, had misidentified.

The first step towards an electronic index involved gathering and unbinding each page of each index for the past seventy-six years. Next, each page was manually scanned using a high-resolution scanner. Since we required high resolution, the scanning process was slow, because higher resolutions require longer scanning times. At the end of this step, we had over 500 scanned pages. At this point, each page was a separate single saved image, not an editable text document. Therefore, in the third step, we used optical character recognition (OCR) to convert the images to text documents. This process must be carried out one page at a time and is more time-consuming than scanning. Also, because we needed to create a single text document
that could be edited with conventional text-manipulating tools, we had to take the
intermediate step of transferring the OCR-generated text files into a format that was
friendly to a standard word-processing program (Microsoft Word). This process was
also not as easy as it might sound, since many times the OCR had difficulties reading
characters here and there, which prohibited us from selecting an entire page at a time.
Rather, for nearly all of the pages, we had to select the text three to four lines at a
time, then paste those lines into the word processor; then return to the OCR document
and select the next few lines, while incorporating any unread characters into the word-
processor file. Additionally, while the OCR software correctly read all formatting and
special characters, many of these were lost on being transferred into the word
processor. The final step, then, and the most time-consuming of all, was to hand-edit
and proofread the entries line-by-line to add missing symbols, characters, and find
any other errors, while also making sure that all entries were in standard Language
index style.

This entire process was not carried out alone. Many people helped with
getting this index into a presentable format, and we owe many thanks to everyone
involved. Special thanks go to our language-specific proofreaders who volunteered
time to edit languages with which they are more familiar than the editors. We are
indebted to our Stony Brook colleagues Tobias Beetz, Christina Bethin, Xochitl de la
Piedad, and Anuntxi Monsalve for their volunteer efforts in editing German, Polish,
Russian, and Spanish. We owe additional thanks to two people who went through the
entire document line-by-line, whose attention to detail should not go unnoticed.
Priscilla Mullins and Zheng Xu spent countless hours editing this index. This task is
both exhausting and tedious, and we are truly thankful to have such patient and
meticulous proofreaders involved in the making of this index. Finally, we would like
to thank Maggie Reynolds for her patience, support, and guidance in all stages of
seeing this index through from planning to production.

The content of this index follow general Language index style with a few
small changes. First, authors’ names are repeated for multiple entries (instead of being
substituted by a dash), because this repetition makes searching the database
electronically possible. Second, we added the volume and year of publication at the
end of each entry, because this is a cumulative rather than annual index. Consistent
with indexes of the past, this index lists the authors and titles of every article
published in Language, along with the authors and titles of all books that were subject
to a book notice or book review. The authors of the book notices and reviews are also
included in the entries.

Our final goal in creating this index was to make it a living document, not
just a twentieth-century index for Language, but one that could grow with the journal
into the future. Because of its electronic format, we can update it, add to it, and make
other changes with sufficient ease. It is our hope that, with the help of the linguistic
community, we will be able to improve this index both by including entries that may
have been missed, or by making changes to existing entries that we may have
overlooked. We welcome any comments that can help improve the index. Please send
comments to Mark Aronoff at Stony Brook, preferably by email
(mark.aronoff@stonybrook.edu). Finally, while we have thanked many people for
their help in making and editing this index, we take full responsibility for any errors
that exist in this document.

This index is the final product of the Stony Brook Language editorial office,
produced by the two people who have had the longest association with that office.
The fact that the index will be a living document gives us heart, because it means that, in some sense, *Language* will always retain a real (if virtual) connection to Stony Brook, for which we are grateful.

Meghan Sumner and Mark Aronoff  
Department of Linguistics  
SBS S-201  
The University at Stony Brook  
Stony Brook, NY 11794-4376  
Email: mark.aronoff@stonybrook.edu