

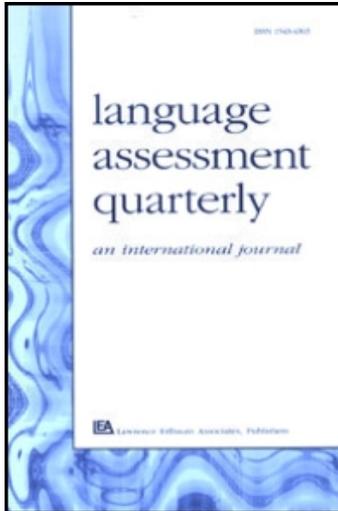
This article was downloaded by: [Saville, Nick]

On: 26 August 2010

Access details: Access Details: [subscription number 791123092]

Publisher Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Language Assessment Quarterly

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t775653669>

From Cloze to Consequences and Beyond: An Interview With Elana Shohamy

Anne Lazaraton^a

^a University of Minnesota,

Online publication date: 10 August 2010

To cite this Article Lazaraton, Anne(2010) 'From Cloze to Consequences and Beyond: An Interview With Elana Shohamy', Language Assessment Quarterly, 7: 3, 255 – 279

To link to this Article: DOI: 10.1080/15434301003792815

URL: <http://dx.doi.org/10.1080/15434301003792815>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

INTERVIEW

From Cloze to Consequences and Beyond: An Interview With Elana Shohamy

Anne Lazaraton

University of Minnesota



Professor Elana Shohamy was born in New York and moved to Israel at a young age. It was an interesting time for the country and the Hebrew language, and Elana shares her keen insights into her experiences growing up in Israel. She obtained her teacher's certificate from David Yelin Teacher's College in Jerusalem, specializing in teaching English as a Second Language. After moving to Minnesota in 1968, she got a B.A. in Economics from Hamline University in Saint Paul and went directly into the Ph.D. program in second language education and measurement & evaluation at the University of Minnesota. Her doctoral dissertation, entitled *Investigation of Concurrent Validity of the Oral Interview Test With the Cloze*

Procedure for Measuring Language Proficiency, was completed in 1978. After 2 years as a post-doctoral researcher at Stanford University and the coordinator of the Hebrew program at UC Berkeley from 1979 to 1981, Elana went to Israel and began her career at Tel Aviv University, where she is currently a professor and Chair of the Language Education Program in the School of Education.

Elana has been a ubiquitous presence in applied linguistics and language assessment, both in print and in person. She is the author of two scholarly books and the coauthor of two more; she has also coedited five additional books, one of which is *Volume 7: Language Testing and Assessment* in the second edition of the *Encyclopedia of Language and Education* (2008, Springer). She has authored or coauthored more than 40 refereed journal articles and more than 50 book chapters. Elana founded the journal *Language Policy* and served as Associate Editor

from 2000 to 2006 and is currently Editor; she has been on the editorial boards of *Language Testing*, *Language Assessment Quarterly*, *Modern Language Journal*, and *Applied Linguistics*, as well as the *International Journal of Bilingualism* and *Critical Inquiry in Language Studies*.

In person, Elana has been an ever-present colleague at language assessment, language policy, and applied linguistics conferences in the multiple roles of plenary speaker, colloquia organizer, panel discussant, and paper presenter. In addition, many scholars and scholars-in-training have benefited from Elana's presence in countries such as Hong Kong, the United States, Canada, Finland, Denmark, Colombia, and Australia in her visiting professorships over the course of her professional career. Elana was a cofounder and co chair of ACROLT, the Israeli language testing group, and served as the president of ILTA in 1999.

Elana's outstanding contributions to the field were recognized on April 16, 2010, when she received the UCLES/ILTA Lifetime Achievement Award at the 32nd Language Testing Research Colloquium (LTRC) in Cambridge, United Kingdom. She delivered the Award Lecture, "From Critique to Change: Building on the Power of Tests," at the last session of the conference. This award recognized "extensive and inspirational service to the language testing community, and/or a valuable scholarship record in the field of language testing." Although she was the seventh awardee, joining a distinguished group of past recipients, Elana is the first woman to receive this award.

Although I had read a number of Elana's publications in the 1980s, I met her for the first time at LTRC in Princeton in 1991. I was absolutely terrified of what she might say about the paper I presented, but there was no need to worry, as she was quite encouraging. The 2001 ALTE Conference in Barcelona was the next time we spent any time together; we did not know each other well, but we took a bus tour to Gaudi's Sagrada Familia and had a meaningful talk at a very sad time in her life; even so, we got in some good and hopeful laughs. The conference dinner in the Hotel Majestic was a turning point for Elana; she has said it was the first time she had smiled in months. Eight years later, Elana visited the University of Minnesota to give two scholarly talks, one on language policy and the other on linguistic landscape. For 3 hours on Friday morning, September 25, 2009, Anne Lazaraton conducted a wide-ranging interview with Elana. A transcription of that interview was produced and was revised numerous times by both Anne (A) and Elana (E).

THE INTERVIEW

A: Elana, I'd like to situate our talk as starting from when you did your dissertation here in Minnesota, so to fill in the blanks, could you tell us about your life up until you came to Minnesota to do your Ph.D.? I know you were born in New York.

E: Maybe what is most relevant here and best relates to language and testing is growing up in Israel in the '50s when Hebrew was still in a process of revival. Especially after the state had been founded in 1948, there was a big flow of immigrants, not only from Europe, as a result of the Holocaust, but also from Middle Eastern countries, North Africa, the U.S., Canada, etcetera. For all these new immigrants the issue of "language identity" was on the top of the agenda, but this included also immigrants like my grandparents who arrived in Palestine even earlier, in 1932. People were "forced" into Hebrew as this was decided by the Zionist movement to be "the language of the land" and a major ideology. Thus, there were informal "tests" administered

to everybody on a moment-to-moment basis, through face-to-face conversation. These were not official tests as we know them today, but every time you gave a public speech or talked to a person in the street, in the stores, in the market, in school, in a public meeting, your language was assessed, “graded” based on how well you knew Hebrew, “the language.” We can say that there was kind of an unofficial CEFR scale, if you will, whereby people got language proficiency scores. I remember conversations at home where we would evaluate the “worth” of people exclusively by their level of proficiency. Hebrew language proficiency was a symbolic tool, indicating belonging, patriotism and loyalty to the idea of the Jewish state; this clearly implied that those with a low level of Hebrew language proficiency, some who had just arrived from Europe or Jews speaking Yiddish, Ladino, or others coming from Arabic-speaking countries speaking Arabic or “Jewish Arabic” or others proficient in a variety of territorial languages, would fail such tests. In my work on the revival of Hebrew I even found evidence of actual semiofficial tests where the level of Hebrew language was also closely monitored by conducting face-to-face oral interview language tests. One of the documents located in the archives, dated June 21, 1939, is an announcement posted in major newspapers in Palestine regarding home visits to assess people’s Hebrew language proficiency. People were asked to cooperate with the testers: “In the town of Raanana, today and tomorrow, pairs of volunteers will visit you in your homes and will conduct a census for the purpose of counting the number of people who know the Hebrew language. You are requested to welcome these couples using good manners.” This was signed by the “cultural committee” of the municipality of the town.

So, how well you spoke Hebrew was a major criterion for belonging. This phenomenon was interesting for me because of my own family. We were living in the same home with my grandparents, and my grandfather was the mayor of the town. So I grew up in a home where two people, my grandfather the mayor and my mother, who spent 7 years in Palestine during the ’30s, were proficient in Hebrew. My mother was born in the Brooklyn, and spoke only Yiddish until age 5, then she acquired English when she started school. But upon moving to Palestine in 1932 when she was a teenager, with very intensive work, including private tutoring in Hebrew, she managed to learn it pretty well. My grandfather spoke Yiddish in New York, but being literate in reading Hebrew texts, did not have great difficulties in acquiring the spoken variety of Hebrew once he moved to Israel. So they continued to speak Yiddish and English at home but could “pass” the “Hebrew identity tests.” It was mostly in public places where it was very important to demonstrate their Hebrew proficiency as they were very aware of the fact that this was how they were being judged and valued. If you spoke good Hebrew without an accent (although it was difficult to decide what the right accent was at the time when the language was new for most people, but there was something of a prestigious accent, Hebrew mixed with Russian, which was more acceptable) that was good. But if you did not speak Hebrew well, you were excluded from the collective identity, no matter *what* you had to say and were often denied employment. So for my grandfather, the mayor of the town, a devoted Zionist but who was a Yiddishist in New York before moving to Palestine and a principal of Yiddish schools, moving to Palestine meant total denial of Yiddish and a need to demonstrate his identity via Hebrew only. But in the same family there were also those who could not pass the “language tests.” So there was my grandmother—Feiga, the wife of the mayor, who was basically at home and continued to speak Yiddish and never acquired Hebrew and like many other women at the time had very limited literacy skills. She also spoke some Yiddishized English from the years she lived in Brooklyn; her Hebrew, even after many years in Israel, boiled down to a very small lexicon in her spoken repertoire. So she

would not go out to many public events, and probably my grandfather felt embarrassed about her lack of Hebrew proficiency. While my mother spent a few years as a teenager in Palestine, my father had never been to Israel before he met my mother in New York in the early '40s. He fell in love with her, didn't care about Israel or Zionism, and just followed her to Israel in the 1950s with their two little daughters. So you have this situation where people like my grandmother or my father were victims, especially my father, who could not find a decent job because of a lack of Hebrew skills. They were both in a way excluded from participation because language was the prime definer of status and inclusion in the new society.

It clearly had an effect on the home dynamics. I remember, for example, coming back from a school trip when I was in the first grade, and my father meets me at the bus, and hugs me and starts speaking English to me. At that moment I totally rejected him because he was speaking English in front of my friends and not Hebrew (which he did not know). So in the '50s it was not only Yiddish that was rejected but also other languages—English (the British had just left Palestine in 1947), Arabic (even for local Arabs who stayed in Israel after the independence as well as for Jews coming from Arab countries), German for sure, and many other languages like Russian or Polish, or Farsi, etcetera. All these were defined as “bad” languages. I would say that even today after many years of immigration and the dominant position of Hebrew, it is still considered central to who you are; English also enters the equation but not as much as Hebrew. So subconsciously, not being aware of that at the time, those formative years in Israel motivated me to go into language studies when I arrived in Minnesota. Language proficiency or the lack of it and its impact on people's lives certainly had an effect on my future interests; I am still identifying with people who are victims of the power of certain languages and the denial of the languages they know.

A: So why did you come to Minnesota?

E: This was “following your husband,” a pretty common trend at the time. . . . He was accepted to study computer science at the University of Minnesota, one of the few places to study this topic in 1969. Right upon arriving in Minnesota I started studying for a B.A. in Economics. I previously received a teaching certificate in Israel in language education, but I was determined not to go into teaching; but still ended up doing it—and clearly my passion was and still is with languages. And as to teaching, my grandfather (Yiddish), my mother (English) and my daughter (cognitive neuropsychology) are all teachers, so I guess it is in the genes. ((laugh)) So when I wanted to continue for an M.A. I somehow found myself walking into Prof. Dale Lange's office in the College of Education at the University of Minnesota and inquiring about enrolling in the M.A. program in language education; “testing” had not even entered my mind at the time.

A: You know, I actually got a copy of your dissertation from the UM Archives.

E: Oh, no! You mean it still exists??

A: I was interested in how you came up with your topic. The Hebrew makes perfect sense to me, but why was it you chose to focus on cloze testing?

E: The Cloze was the “hottest” topic at the time; it was “sold” as a magic tool, kind of a trick for testing. It was argued that it could be easily constructed and graded, it was cheap and efficient, and it could be used to test “everything.” I read John Oller's publications and started to correspond with him, showing my interest in cloze testing. He was interested in people doing research on cloze because it was correlated successfully with all skills (we were still in the skill era then . . .) except with spoken language. I was not interested in testing at all; in fact testing was not “a field” at the time! ((laugh)) Then Dale Lange had visitors from ETS or the

CIA who came to the university; they were doing the FSI interview in all government offices at the time, but mostly at the FSI, CIA, and Peace Corps. But universities were not involved, and there was no research on that type of test anywhere at the time; and remember that the test was innovative in that era as it was the only test where candidates actually spoke face-to-face, rather than repeating audio sentences they heard on a tape recorder. So the U.S. government agencies had an interest in “spreading the word” and especially in finding out about the cloze’s psychometric qualities—reliability and validity. I think they also thought that if universities got involved, then these tests would gain legitimacy and be of even greater use at universities and secondary schools. They basically said, “Nobody is doing speaking tests where people actually speak and we have lots of experience in this, but we are not sure whether it actually works psychometrically,” or they may have thought that the tests were really good, but they needed the credibility from academic institutions . . . in one conference I remember them saying that the oral interview has “emotional validity,” or “experiential validity.” So these government people arrived at the University of Minnesota hoping to find graduate students who would do their dissertations on the FSI oral interview. Obviously we knew very little about the test as there were hardly any publications about it at the time.

They were mostly interested in Ray Clifford and Jerry Larson among the five doctoral students studying with Dale. We were a very active group of students, teaching methodology courses in the department every Monday afternoon with Helen Jorstad and Dale, who were our leaders; we were speakers of a variety of languages and I worked with the Hebrew teachers. So we got very good experience and training in language learning theories and teaching, but nothing about language testing; I mean there was hardly any research, there was no organized group of scholars interested in it, etcetera. So when the government people showed up at the university and proposed we do research on language testing, it was very appealing, something new. In other words, to answer your question more directly, with the interest in cloze at the time and then the exposure to the Oral Interview, somehow the combination of the two went well together. So they took us to ETS and on the way we stopped at ACTFL in Washington, D.C. in 1975 where John Clark presented a very impressive paper on direct versus indirect tests. For the first time I understood that actual face-to-face interaction for oral testing was the only way to go; it really affected my work on testing for all these years, the idea that language has to be performed on tests. I clearly remember the continuum he presented showing the differences between the indirect and the direct ones which included extended interactive speech. So from there we continued to ETS in Princeton and worked with Pardee Lowe and John Clark to acquaint ourselves with the FSI oral interview, get some training on how to administer it, assign scores, and use the rating scales they developed, which were very innovative at the time. I had not seen such a rating scale before that actually defined language proficiency in a hierarchical manner according to different dimensions; until then people just counted errors, so that was also very new. They hoped we would all do our dissertations on the oral interview. So when we returned to Minnesota, we did that. This is what got me going.

I never planned to go into language testing. I find that this is the story of many women, not having specific plans or goals but taking advantage of opportunities and when you do, you almost always benefit. This is very different than men of that era. As a woman I had no plans, no specific career goals, but I was lucky to have opportunities (or maybe I pursued them too . . .); in most cases it turned out that these led to something good. I often think that not having specific goals is healthier as it takes you to unknown, new, and exciting places, and makes you appreciate more what you have; perhaps those who do have plans, when they do not reach them or they fail to

materialize, they feel disappointment. I believe that women of today are to some extent different as they are more geared to their future at an earlier age, maybe not to specific goals, but at least they know that doing something interesting at work is so valuable and that it is crucial to one's identity and happiness. But I often think of other women of my generation and especially older women who missed opportunities, and what society missed by not having them doing things which are more meaningful. I am very encouraged by the younger generation of women who have different views. I am very grateful for the interesting life I am having, doing exciting work, pursuing new ideas, expressing my views publicly, finding people who are interested in what I have to say, being invited to so many places in the world and learning about language situations in places such as the Baltic countries or China or places in Europe. I am grateful for working closely with curious students who read and research these topics and who see the central role that languages play in the daily lives of people, especially for those who are excluded because of language; they are taking language studies to a very interesting place. So I am happy to see that more and more women have active academic careers whereby the work they do can contribute to this world.

So as to language testing, the rest is history; the others doctoral students and I did our dissertations on various dimensions of the oral interview. In my own study, I interviewed 106 people doing the oral interview in Hebrew who also were tested by the cloze in Hebrew in various formats. Oller was very encouraging as he wanted to see correlations between the speaking and the cloze, which turned out to be high. And then we had the first LTRC in Boston in 1979 where we just created a language testing group out of a small meeting. I got there via an invitation from Oller. Lyle Bachman was there. Charlie Stansfield was there. Buzz Palmer was there, and so was John Clark. I remember I met Andrew Cohen for the first time, who started recruiting me to go to Israel, because I hadn't thought of going back there at the time. Merrill Swain and Michael Canale presented their now famous Canale and Swain model and a number of us presented our dissertation results, and attention was given to the oral interview publicly. In retrospect I must admit that I am not that fond of my dissertation; I think the kind of work and research I am doing today is more human, less technical, and it contributes more to justice and equality and is more attentive to people—it has more social implications, far more than that dissertation. I am very glad that I detached myself from that work and went in a different direction in testing, with a focus on its social and political implications.

So there were LTRCs in Boston, San Francisco, and Ann Arbor, and I went to all three. It's now been 30 years that we have been meeting annually to discuss various aspects of language testing. At the time, it was a good feeling to find a group of people with mutual interests. I find this is a very good strategy in the profession, whenever you go into some new area you need to find a community with whom you share your interests. They provide legitimacy for your interests, which is often lacking in groups that you have spent some time with. It is like a family that you need to detach from to gain new identities.

A: You know, I didn't read your dissertation. I looked at it because it's a historical document—you were doing something sort of cutting edge at that point. One thing I noticed was number crunching—that's what you did at that time was number crunching, right?

E: But it was simplistic, only correlations and very simple regressions; people were doing much more interesting work right after that, the multimethod multitrait, the different factor analyses, etcetera; my work was kind of in the early days of that era.

A: Do you consider yourself a stats expert?

E: ((laugh)) At that time, the only way to be in this profession was through statistics (it is like gaining identity via Hebrew in Israel . . .). So when I was studying for my Ph.D. in Minnesota I was majoring in Education, in the program of Second Languages and Cultures and my minor was in Psych Foundations. So I had many courses in statistics and measurement and I also specialized in program evaluation with Michael Patton from the Sociology department. He was a big name in evaluation at the time and he directed students to examine evaluation utilization and consequences. So, when I think about it now, he was probably the first one who planted in my brain the notion of test consequences, impact and use; that is, the utility of results of tests and evaluations for bettering society and education. When I took his courses I started thinking of tests in that way as well—“Okay, so we created a good test that is reliable and valid, but ‘so what?’ What will its impact be? How will it better this world?” Program evaluation and especially evaluation utilization was about “let’s do something with the test and its results, let’s examine what it really does; otherwise, why bother?” Evaluation was relatively new, but it had already developed solid theories and designs, because it emerged from the work in measurement and research.

A: In your dissertation your conclusion was that the cloze had fairly good correlations with the FSI. It was efficient and avoided the hassle of doing a thirty-minute interview with raters—it’s sort of an acceptable solution to that problem. How do feel about cloze testing today?

E: That was one conclusion, the statistical one, but I was not convinced by the results, because already by then I did not buy the implications of statistics in that way. Thus, in spite of the high correlations between the two types of tests, I could not accept statistical predictions—they were too abstract and did not make sense—I never fully trust statistics. When it comes to language I could not accept that any construct that does not involve interacting with actual “real life” language with actual people could provide a meaningful answer. I was already convinced that there may be other intervening variables that create the correlations, so when you see language use in action, whether in writing or in speaking, you see actual production. Language in use is the only criterion I trust, even with all its instabilities: It may be complex but it is real, language *is* complex and there are no shortcuts. Then I started asking questions about the uses of the test, because questions about washback were important with regards to utilization; so if you cannot convince language testers as to the advantage of direct tests versus those that predict well, maybe you can convince them that direct tests have better effects on school and learning and teaching. So I didn’t go in the direction correlations suggested. I was not interested in prediction but was very interested in behavior, actual language use, in people acknowledging the complexities of language rather than running away from them, which testers generally do. For all kinds of reasons testers are constantly looking for simple solutions, such as the CEFR, for much more complex issues.

In addition, it was around that time in 1982 that I went back to Israel, where oral tests had been administered in English as a foreign language as part of the national matriculation examinations around the country to all students of 12th grade for many years. The oral test was actually a “conversation,” an interview between a tester, who was normally a teacher, and a student. When I met Thea Reves she was working as an English supervisor in Israel and she experimented with a “group interview” where four students got together and discussed a topic of interest that had been assigned to them. But they did not have the statistics, interrater reliability, concurrent validity, and so on.

A: What you’re saying is that this was back in 1982, so the current interest in group orals is not new?

E: Exactly. I remember well when I met you, Anne, at LTRC in Princeton in 1991, even then nobody was doing this kind of work. So when I had just arrived in Israel the idea turned me on. The situation is very different than in the U.S., because in Israel, where the educational system is national and centralized, you can make a decision or a recommendation to the national supervisor or the Ministry and if they accept it, it becomes policy and affects all students in the country immediately. I only then realized the power and impact of tests but also the responsibility of testers. So once in Israel, Thea Reves, Yael Bejarano, and I decided to run an experiment with the “group test” and the “oral interview,” and we added to the battery a “role-play” and a “reporting” task. We invited hundreds of 12th-grade students to Tel Aviv University on a Friday morning and each student took *all* four types of oral tests. They switched from the group test to the oral interview to the reporting task to the group discussion, with different testers, in different rooms and in mixed order, and we then examined each of the interactions in terms of reliability and validity. The main point is that it was very obvious that the student produced oral language differently depending on the type of interaction. Thus, some of the students had high proficiency in the oral interview but found it difficult to express themselves in the group discussion; others were fluent in the role play but almost silent on the reporting task, and so on. Based on the correlations and different statistical analyses we were then able to convince the English Inspector at the Ministry of Education to introduce an oral test battery that included *all* four types of interactions. Indeed within a few months this became the new national policy. Clearly, the process of policy making in Israel was different, for the good and for the bad.

A: Is it still policy now?

E: In a way it is—but not the very same assessment battery because over the years it went through a number of transformations. At the time we tried to convince the Ministry that what worked as an experiment would not necessarily work in the same way when it is applied “on the ground.” It was surprising for me and for my colleagues that our recommendations became policy with no pilot testing; I had many nights with no sleep, feeling the burden of responsibility.

A: So were you responsible for those tests for some period after that—did you have oversight of them?

E: That is an important question. This was our recommendation for a number of reasons; the Ministry accepted it for *their* own reasons and they were not interested in following up on the new testing policy once it was implemented. As a tester I always come back to this question—what is our real responsibility as testers? Are we responsible for recommending policies based on our small-scale research studies? Or is it our responsibility to follow the implementation of our recommendations as they are applied “on the ground,” in “real life”? I learned the hard way in the past few years, especially with regards to a study I did on the length of time it takes immigrants to perform academically in schools in relation to native speakers. The answer is on average 9 to 11 years, even for immigrants coming from the former USSR, who arrive with very high level knowledge in math, for example. But the lack of knowledge in Hebrew slows them down in math, which is taught in Hebrew only. So is it our responsibility once we submit the research report to the government to also follow the consequences and the policies which are based on these recommendations. In fact, some of the policies that emerged from these recommendations were to separate classes or to “bus” students, while we never even hinted at that.

So, what are the real responsibilities for us as researchers when we do end up with recommendations? This is a somewhat similar case to the language education policy which Bernard Spolsky and I recommended to the Ministry of Education a decade later. In that case our recommendations

indeed led to a new educational policy for Israel, but it was only a document that did not lead to any meaningful implementation, and clearly no studies examined the extent to which these recommendations were actually followed. They were in a hurry to create the policy document but in no hurry to implement it or to examine whether it was implemented at all. But back to the oral testing battery at the end of 12th-grade English, students nowadays perform in a variety of oral interactions, interviews, and role plays, but they also do more creative things such as projects and portfolios, but not based on any research.

For me the whole experience with the introduction of the oral tests then led me directly to the work on the power of tests. I realized that the Ministry of Education used tests as policy tools to advance very different agendas, specifically as power tools over teachers. Now that they had this oral battery, it could be used as a whip to force teachers to teach oral language in the classroom, something that teachers had not complied with when other devices were used, such as a new curriculum. So the new oral battery forced teachers to teach speaking and students to practice oral language in preparation for these very high-stake tests. I could see then with my own eyes how in centralized educational systems that teachers are not independent thinkers but rather people who follow orders that are imposed on them on a regular basis, and these are almost always given in the form of tests; these tests determine the agenda, and the Ministry is very aware of this phenomenon and takes advantage of it. Tests are ideological tools that determine policies and guarantee compliance; the main point was to use the tests as a mechanism to force a change in the teachers' bad habits.

A: So they wanted your credibility.

E: Oh, yes! But once I became aware of the unquestionable power of tests I initiated a study that examined not only the effect of that oral test battery on learning of spoken language in the classroom but also of two other major tests, a national test on reading comprehension and a test on the Arabic alphabet. In all cases, we found that the teaching was significantly affected by the format of the tests; in the case of the oral test, it was clear that the introduction of these tests resulted in teaching "oral test language," as only those very oral interactions used on the tests were taught in the classes.

A: You mentioned the 1996 revision of language policy in Israel. What was changed?

E: It is not exactly a revision as there had not been an overall language education policy in Israel before. I spent some time at the National Foreign Language Center¹ in Washington, D.C. in the '90s and was exposed to the topic of language policy as they were trying to introduce language policies in the U.S. At the time there were visiting scholars from different places who shared the language policies of various countries in the world. I presented a paper for the NFLC about the language policy of Israel, arguing that the policy was indeed "Hebrew only" and was totally driven by ideology, that is, the revival of Hebrew as a language of Israel, and that along the way, this aggressive ideology led to the suppression and eventual disappearance of the large number of languages that were spoken in Israel by immigrants. So the main point of the paper was that when language policy is driven by national ideology, there is a high cost so that diversity disappears in the name of language and for the sake of the idea that one language is needed for cohesion and unity. Bernard Spolsky spent some time at the center as well, and there was a time when we overlapped; we both got very interested in the topic of language policy and spent long hours

¹National Foreign Language Center: <http://www.nflc.org/>

discussing it. So when we both got back to Israel I sent that paper to the Minister of Education, and within a few days he approached me to ask about issues of language policy and requested that Spolsky and I write a document that would propose a language education policy for Israel. It was pretty shocking how fast these things moved, and even when we suggested doing some research first to find out more about what is going in the schools, he feared that he would not be in office very long, asked us to do it within six weeks, so we did. We recommended a policy of *Three Plus*, which meant that Jewish Hebrew-speaking children would have to learn Hebrew, English, and Arabic. The “plus” implied the need to learn additional languages such as community, heritage, immigrant, or world languages; we strongly encouraged maintenance of immigrant languages. Students born into Arabic would learn Arabic, Hebrew, and English, plus additional languages. Indeed after a few meetings, our proposals became the official language education policy for the country. It is important to emphasize that this was not the national language policy for Israel, which states that Hebrew and Arabic are the official languages, but this refers to the language policy to be implemented in schools. Within a few months, the Minister of Education, Ariel Rubinstein, and the Director General, Shimshon Shoshani, were removed from their positions as the government was changed; still we were funded by them for three years to examine the educational policies which were being carried out in schools. In the 1999 book *The Languages of Israel*² that I wrote with Spolsky we summarized the work we did following the proposal and the implementation of the policy and we examined the de facto education policy of Israel. Yet, since 1996, when the document became official policy, nobody has examined its actual implementation; today we know that besides some increase in the learning of Arabic among Hebrew-speaking Jews, not much has changed. For various political reasons nobody really paid much attention to the new policy. There is no need to elaborate on the complexity and contested issues in Israel but it was clear that without activism and real financial and educational support, declared policies do not have much impact and they often remain just words on paper. The lack of interest by the new government in these language issues certainly contributed to the lack of impact of the policy.

A: So the policy just goes on?

E: The document is still there, but it is clear that policy and deeds do not always match. Currently there is pressure “from below” to create different policies, that is, learning English from an earlier age, studying Chinese as well as Spanish, including spoken Arabic within the study of Arabic so that students will acquire communication skills, and even some efforts to learn heritage languages such as Yiddish, and strengthen the study of Russian; these have resulted in some pockets of change here and there. Recently there has been some interest by the current Director General of the Ministry of Education, who happens to be the same person who was there at that time, to rewrite the policy. We will see what happens.

A: Do you speak Arabic?

E: No, not really. In fact I view myself as a victim of the lack of language education policy when I was a student. At the time there was no overall policy, but students would mostly learn French or could choose between French and Arabic and I chose French. Arabic was not very popular; in fact most students did not learn any other language in addition to English. But this

²Spolsky, B., & Shohamy, E. (1999). *The languages of Israel: Policy, ideology, and practice*. Clevedon, UK: Multilingual Matters.

is changing now as there is the realization that living in the Middle East where Arabic is the major language and is used by Arabs living in Israel, makes it a very necessary language. As I mentioned before, Hebrew swallowed up many languages that Jews spoke, like Yiddish and other territorial languages; these have vanished from the linguistic repertoire of Israel. So I've been trying to draw people's attention to the high cost of reviving Hebrew and the importance of maintaining the languages that immigrants bring with them to make Israel a more multilingual and plurilingual country, as is the situation now in Europe.

A: You've told me about the Three Plus educational policy in Israel today. What kind of assessments go along with the policy? Are there exit tests after high school? Are there tests for each of those three languages?

E: This is indeed a major issue as the tests do not match the language policy at all. All the tests used at the end of secondary school or as entrance exams to higher education involve only two languages—Hebrew and English. Students also take entrance tests to higher education, similar to the SAT, which are given in multiple languages, but the Hebrew tests are the primary requirements for entering Israeli universities. This is for all prospective students, including those who learned Hebrew in school via Arabic. So, even if some of the languages that I just mentioned are learned in schools or are part of the written policy, it has not changed the testing policy at all—very few students study the languages for more than 3 years, and very few take matriculations exams in these languages. All students are required to take Hebrew and English graduation tests, and Arabs are being tested in Arabic, Hebrew, and English, as Arabs are trilingual. But for Arabs who attend Israeli universities, all classes are taught via Hebrew as the main language of instruction. Thus, Arabs are losing their motivation to maintain their Arabic, which is the language of instruction until they reach high school. Thus in many of these schools and in many subjects, Hebrew is penetrating the Arabic educational system as they realize that in higher education it is only Hebrew that has status, along with English.

A: What are those tests like?

E: Well, let me say before anything else: the Israeli educational system is dominated by tests. Just as an example, a few years ago, there was a big drive to create an educational reform—nothing really came out of this in spite of huge financial investment—which was to establish a new testing agency as part of the Ministry of Education. It was agreed that it would be as an external body that would also examine the functionality of the educational system. The testing agency produces or subcontracts many tests nowadays and Israeli students are clearly overtested by the Israeli version of No Child Left Behind and various international assessments such as PISA³ and PEARLS.⁴

A: I'm not very familiar with PISA. Can you describe it for us?

E: PISA is an international test, part of the OECD,⁵ that is running in more than 50 countries. It is the biggest thing around now with a huge influence on educational systems around the world, especially in Europe. PISA is very problematic as it is about the world becoming standardized and homogenized. Those tests—the PISA and other international tests along with the

³Programme for International Student Assessment: http://www.pisa.oecd.org/pages/0,2987,en_32252351_32235731_1_1_1_1_1,00.html

⁴The Predictive Early Assessment of Reading and Literacy.

⁵Organization for Economic Co-operation and Development: http://www.oecd.org/home/0,3305,en_2649_201185_1_1_1_1_1,00.html

Israeli MEITZAV⁶—now serve as the main criteria for examining the quality of education and for defining knowledge. Israel was on a very low ranking on the PISA, so rumor has it that the new Director General was nominated for the job on the basis of his guarantee to bring Israel to a higher rank. It is also the case that classes taking these tests get awarded with special budgets, small class sizes, and quality teaching, so they can do better on these tests. Thus, these international tests become *the* measuring stick for the standing of nations in the world and for judging the quality of education. The matriculation examinations at the end of secondary school are not as important as the PISA and the other tests, which became the tools to judge the status of countries in the world. Think of the far-reaching economic implications. What about the local knowledge that gets lost along the way, the students who are taught to succeed on these global tests so they can serve national ideologies, and the intensive preparation for these tests that occupy the educational priorities of so many schools??

A: This reminds me that in your 2001 book on *The Power of Tests*,⁷ you wrote that “for a number of years, I felt that my colleagues in language testing did not appreciate my ideas concerning the use and consequences of tests.” Has that changed?

E: I think it has to a large extent, because when I first presented these ideas I felt that language testers rejected them. I felt that the people working in this area of language testing viewed me as a betrayer of the field, that I was doubting their identity, their ethicality; I felt they took it very personally. If we look backwards at the history of the field, language testing was originally based on psychometric theories and in my view was very technical. Many of the participants in the field had an interest in test development, which I find to be less academic and less deep. And today as well, I find that the papers presented at our conferences are very technical and fail to address larger questions about implications. I find that we generally try to address questions posed by testing organizations such as ETS or Cambridge ESOL. I find the researchers in the field are not interested in the effect of these tests on human beings, on the cost of tests, on what language means in this day and age of immigration, globalization and transnationalism, and on various societal dimensions of tests. I find that year by year I am becoming less and less interested in the field; when I compare my interest to the papers presented at AAAL, these are broader, deeper, and have larger implications. I simply do not see an intellectual challenge and I find that our tests are detached from people.

For example, very few people were interested in our own symposium at LTRC Denver last year on language tests for citizenship, which are sweeping Europe. It seems to make them very uncomfortable talking about these issues. They refuse to be critical—I think they want to do the work of doing tests but refuse to look at consequences, somehow being afraid of them rather than facing them. Like what if their tests lead to deportation of immigrants? I find that testers, still today, would rather stop at the level of providing a test and not ask these hard questions, in a way like how I provided the Ministry with the oral tests in the past. But the field has developed—at that time we did not think that our tests had consequences but now we do, so can we treat tests in

⁶The “Meitzav” exams, a Hebrew acronym for school growth and efficiency indicators, are standardized tests that are given to pupils countrywide at the end of the each school year. They are meant to specify just how well Israeli junior high and high school pupils are meeting the benchmarks and standards specified in school curriculums: http://cms.education.gov.il/EducationCMS/Units/Mazkirut_Pedagogit/English/Meitzav_Info/Meitzav_Examinations.htm

⁷Shohamy, E. (2001). *The power of tests: A critical perspective on the uses of language tests*. Singapore: Longman.

the same way? Yet, I feel my work is appreciated, or why else did I get invited to Berlin,⁸ where language tests for citizenship were discussed at the Goethe Institute? They knew very well what I was going to say, that I would be against these tests, so why did it turn out that they published my paper as one of the very few in the conference proceedings in a general publication? Possibly it is because there are agencies and organizations that want to show that there are more voices to be heard and that the field of language testing is democratic and inclusive and opposition is welcome as well. Maybe it is because it is nice to have a critical voice as it transmits a message that “We have morality, we have ethicality, we think about those things, about consequences and people.” So I would say that whereas in the past, my work did not get any attention from testers while it only received attention from applied linguists, nowadays I am very much included and recognized and even encouraged to voice my views about opposition to certain types of tests that have detrimental effects on people and are used as deportation tools.

But does it mean that there is anything different in terms of the use of tests? I am not sure—maybe, maybe not. It may be good for governments to hear the opposition or talk to policymakers about the negative ramifications of these tests; maybe they will use those tests but with more caution. But for me it is still encouraging to see that there is a critical mass in this field today who pays attention to these topics, researches and writes about them, teaches critical testing topics in their courses, and give[s] seminars to young students about them, so that now the various areas within critical language testing are integral parts of the discipline. This is far more than what we see in other testing and measurement groups, more than what we see in NCME,⁹ for example. I think that we have been influential as there is ample talk about consequences, washback, morality, policies, codes of practice. Whether testing agencies are actually applying these concepts in test development and use is another question—that is, do they discard tests with negative consequences? We do not have much information about that, and I wish we would have more papers and research on that issue. Perhaps, I often think, other researchers need to do this kind of research, to monitor testing agencies so the tests with negative consequences will not be used, and to protect test-takers from such tests, by introducing issues of test-taker rights, which should be part of civil and just societies.

For a long time and still today, tests have been used with no monitoring whatsoever and their power is rarely challenged. Thus, those developing the language citizenship tests will not do the research as they benefit economically from these tests, but more independent researchers at universities who do not benefit financially from these tests should take responsibility for carrying out such research and reporting the results in widely read publications. This is like what we do with drugs and medications, imposing penalties on those who continue to use substances which were documented to have negative consequences. I think our field must draw awareness and attention to the power of tests, their dangers, and their problems. We certainly need to continue with research along these lines, to examine test consequences, good and bad, and to interview people who were victims of tests and then publicize these findings. I think that the special issue of *Language Assessment Quarterly*¹⁰ was important, but it needs to reach a wider audience so the

⁸Language and Integration Conference, Berlin, September 18–19, 2009.

⁹National Council on Measurement in Education: <http://www.ncme.org/>

¹⁰Shohamy, E., & McNamara, T. (Eds.). (2009). Language assessment for immigration, citizenship, and asylum [Special issue]. *Language Assessment Quarterly*, 6(1).

public and various civil rights organizations are aware of these dilemmas and consequences; it is our obligation to point these things out and act for changing injustices when they are detected.

A: In your dissertation, you quoted Spolsky, who in 1978 was talking about testing having a prescientific era, when it was just teachers doing tests. Then there was the psychometric era, which to some extent still exists. And then, according to Spolsky, there was the integrative sociolinguistic era. What era are we in now? And, by the way, who is “we”?

E: I think that we are now in an era of uncertainty. In the *Encyclopedia of Language and Education* (Vol. 7)¹¹ which I edited, I wrote that the concerns of language testers in the past decade about the use of tests and their political, social, educational, and ethical dimensions have made the field more complex and uncertain, meaning that new questions are being raised about the meaning of language and the possibilities of measuring these complex and dynamic variables. It is also an era of an ever more compelling need to ensure that tests are valid. But validity also means the protection and guarding of the personal rights of test-takers as well as positive wash-back on learning by addressing the diverse communities in which the test are being used. Thus, the emphasis is also on how these tests can ensure that we will be more inclusive, democratic, just, open, fair and equal.

As to the “we”? I think there are a number of us in the profession who ascribe to such views by addressing tests as *tools*; other use tests as *targets* and they are still seeking the perfect test. But the perfect test psychometrically may not meet the other criteria that I just mentioned. For me, I want to see how the tests we are developing are part of a social and economic structure, and encompass language policy issues where tests are being used by governments, defining nations, and selecting people. So I think for me, and for others working within this paradigm, tests are weapons that need to be watched, followed and monitored.

A: But that was what people were thinking—tests as targets—that’s what you did.

E: I did do that in the past; I have no regrets, but things have changed, language and language tests are complex now, more political, in a world with vast immigration and globalization, where language tests are being used in ways we have not seen before. At the same time we are realizing it is a more multilingual world, a place where nations attempt to control people but people refuse to be controlled and they continue to use languages or live in their own way. I mentioned earlier the research on the length of time it takes immigrants to acquire academic language and noted that it takes about 10 years; given these findings, how can we give immigrant students tests after they have been in a country for less than a few years? It is obvious that they will fail these tests and will get stigmatized because of these results, even in their own eyes; should we not pose these questions to Ministries of Education about the unjust policy of testing these students? Should we not monitor these policies and fight against them to protect these students? In the same context, take for example the CEFR, which focuses on one language only, a monolingual ideology of “national” language. The CEFR overlooks the reality that immigrants develop a mixed bilingual code that stays with them all their lives. It is as if tests try to educate people about how they should use language rather than reflecting the kind of language and varieties they actually do use. I object to tests as ideological tools when they attempt to change behaviors rather than reflect it.

¹¹Shohamy, E., & Hornberger, N.H. (Eds.). (2008). *Encyclopedia of language and education, Volume 7: Language testing and assessment* (2nd ed.). New York: Springer.

A: I'm wondering how these ideas take shape in teaching about language assessment. For your talk today, "Ideology, Practice, and Mechanisms in the Establishment of Just and Ethical Language Policies,"¹² I am requiring my ESL methods students to attend. I asked a colleague whether students in language assessment classes would be required to attend, and this person said, "No, it's not a talk about language assessment." What's your reaction to that? Can you describe a language assessment class you have recently taught?

E: ((heavy sigh)) On the one hand, we learn how to design various assessment procedures; it is not very technical though as we start with the big questions about the complexities of language, about how people acquire language in different contexts, about how language varies from one situation to another, and it is not uniform or homogenous. We learn that there is no one type of language knowledge and what is good for one person may not be good for another, and that language in today's world is different than the closed system we used to think it was in the past. We therefore operate under the assumptions that there are not homogenous answers to testing and assessment, that there are serious consequences of tests, that they send strong messages to students and to teachers. We have to develop a sense of responsibility about the procedure we are learning about, as we are playing with fire. Yet, we do not work in a vacuum, and it is important that while we are trying to change and improve and adjust our tests, we remember that we are dealing with the real world in which our test-takers live, a world where teachers cannot ignore the pressure they get from the outside world for their students to succeed on these big tests. But at the same time my students as teachers have lots of areas they can challenge and protest and negotiate, and they can become activists for changing realities, certainly in the classroom, but not only. They also learn that every time you administer a test you should ask the "why" question: Why are these tests being given? What are the intentions of policymakers when they introduce tests? The taking of a testing course professionalizes them and grants them a voice so they will be listened to.

But it is strange that in the LTRC group these questions are often not posed. I remember even in LTRC in Ottawa in 2005 we gave the final talk about the future of language testing and afterwards someone came up to me asking, "Why do you hate us, Elana?" I often find that in this profession you have to be patriotic and this often means not asking the hard questions. It is the same as living in Israel, where people often interpret my political views with, "If you do not agree with the national policies, why do you live here?" I think that working in an area such as language testing and playing with such powerful tools as tests, and living in Israel, which is such an ideological and contested place, you have to constantly ask the big questions.

A: I found this book of yours, *A Practical Handbook in Language Testing for the Second Language Teacher*.¹³ It really is a practical book. Since the time that it was written, did you ever think about publishing it?

E: Yeah, many times in fact. I tried Oxford and they thought it was not theoretical enough. Every few years somebody comes up to me and says, "Let's publish it in some way; I have been using it for so many years, and it is so useful." It is now 25 years old, and clearly major changes have taken place in assessment design, so I wish someone else would write a book that is teacher friendly, as most books are not like that. I think it is popular because it talks directly to and with teachers. I think it is the kind of book that should be put on the Web for everyone to use and

¹²Talk given at the University of Minnesota, September 24, 2009.

¹³Shohamy, E. (1985). *A practical handbook in language testing for the second language teacher*. Ramat Aviv, Israel: Shaked.

teachers (and maybe even students) should add new items and tasks and ideas, kind of like a testing Facebook.

A: What is your opinion about dynamic assessment?

E: In my view the principle is very right because it is blurring the “testing” and “teaching” divide and puts the emphasis on real learning rather than on judgment. The research so far is mostly pointing to specific cases where it works but I would like to see more large-scale studies. I think the theory is very powerful, and therefore we need to find creative ways to use it more widely. One of my doctoral students is currently writing her dissertation on the use of dynamic assessment in the context of end of secondary school exit tests using speaking, as we just talked about. If the results show that it is effective for introducing mediation as part of large-scale assessment and that students actually show progress on their scores as a result, that will be an indication that we can introduce it on large-scale oral language tests of English. Clearly doing large-scale assessment that can incorporate learning is the ideal way. I think it is really powerful when theories from language and learning enter the field of language testing, I wish we would see more of these. It is also very important to mention the huge progress made in the past decade in classroom assessment, or what Chris Davison and Constant Leung refer to as teacher-based assessment.¹⁴ Their work along with Pauline Rea-Dickins and Liz Hamp-Lyons’s ideas about developing practices, theory, and research in school and classroom assessment (some employing dynamic assessment) are notable examples. I am very interested in ways of connecting large-scale assessment with learning whereby it is not that large-scale assessment feeds into classroom testing but the other way around.

A: That’s interesting. Can you say something more about assessment-for-learning?

E: I would say that it is a response to washback, because once we discovered that tests have an effect on learning because of their power and their impact on decision making, we realized we cannot separate assessment and learning anymore; they are connected. So one contribution of the washback studies is that testing and learning *are* connected but in a negative way, as it was shown as a one-way flow as to affect teaching, but what about the other direction, does teaching affect testing? We are still giving too much power to the large-scale tests to affect learning. Except now, in testing *for* learning we are demanding that we need to understand what tests actually do and if they do not lead to learning, as we hope the dynamic tests do, then we should not use them; this is a major challenge for language testers. I think it is very interesting to watch the current debate about the NCLB as it is being reauthorized by the Obama administration and find out that these penalties and sanctions do not work and that just making a decision that on a specific date, say 2014, we will be able to “close the achievement gap,” just because we are testing, and testing is an illusion, a false hope. Likewise to think that by imposing the PISA test and teaching for that test we will be able to close the gap between rich and poor countries is false as well; perhaps it reflects the ability of students to solve test questions, and that is a very narrow criterion to judge achievement in this day and age. I think it also defies the notion of classroom learning and introduces the homogenous definition of global or national knowledge into the classroom while overlooking the diversity of society, towns, neighborhoods, backgrounds, and contextualized knowledge.

¹⁴Davison, C., & Leung, C. (2009). Teacher-based assessment: An international perspective on theory and practice [Special topic issue]. *TESOL Quarterly*, 43(3).

A: What you've said just now reminds me to ask: Are you a good testtaker?

E: ((laugh)) No, I'm not a good testtaker, I'm not a good language learner. This is why I identify with victims. I actually do not like tests at all; this is what gives me the ability to observe, to gain insight into the testing process as I come to it from the perspective of the victim. Something that occurred to me the other day—I was teaching my granddaughter, Alma, who is 4 and a half, the concept of “riddles.” I would ask her a question and she needed to give the answer, like—I would ask her—“What is outside green, inside red and has black seeds?”; simple right? I could see how the question really annoyed her. I could see that she was thinking, “Why would she ask me questions that she knows the answers to?” And then I thought, it annoys her because it's like a test. Here we are having nice conversations, I am explaining to her things about life, history, geography, and in fact she keeps asking me questions, clarifications about what we talk about, “why is this and why is that,” and then . . . I switch into a “testing mode” where I am asking her to provide me answers for things I clearly know and she may not. At the end she may have enjoyed the fun of it, the tricking, the game, but basically it's not something that is very natural. I hate it when very knowledgeable people, often men who are proficient in statistical facts, ask difficult questions that expose the ignorance of those being asked in relation to them—like how many people live in Antarctica. It is often very trivial information that could be found on the Internet or in encyclopedia if you only bother to look. I always perceive it as a power game. With my granddaughter I could see she felt uncomfortable about it. I think testing is an artificial type of genre; performance is fine, projects, OK. So I wish we could get away from the classic testing genre of the “knower” addressing the “non-knower” who has to prove he or she “knows” what the “knower knows” via all kind of tricks, like choices, or distractors that only aim at even further confusing the one who is being asked. I find it creates fear, anxiety, stress, and distrust and even some type of indoctrination.

A: I'd like to switch gears here and ask you about something we've talked about in the past and you alluded to earlier. What do you see as the gender gap in language testing? The consensus seems to be that in language teaching in general, the academics are the men and the practitioners are the women.

E. Certainly, it's a boys club; I noticed it from day one. In most cases women do the teaching, but it also depends on the field and it is fast changing; in Israel most school principals are now women, for example. I think that in spite of major gains for women, some areas are still dominated by men as women are in the process of moving in. In language testing, still the “old” guard of my generation, the majority are men. But the younger generation consists of a growing number of women, and I am just amazed how in my years in the profession, things have changed so fast, as I clearly remember being one of the few women around, along with Liz and Frances Hinofotis, in the early days in the field. Maybe we should do a paper on that at LTRC about the history of LTRC from a gender perspective. I think men just do not understand what women were going through being so marginalized. I certainly wish that women in language testing would become more active, as we were too silent and complied too much. I often felt it was not the right place to be, and I would go to other tangential areas in applied linguistics; possibly the move to the critical dimensions may have been part of it. I feel closer in the past couple of years, especially working on the encyclopedia and realizing that the work is more appreciated.

A: You've done quite a bit of productive work with Bernard Spolsky. When did you first meet him?

E: I never studied with Bernard, but we have been working together a lot and keep continuous contact, almost daily. I clearly remember when I was deliberating whether to go back to Israel in 1981, and it was very motivating to know that someone like Bernard, whose work I so much appreciated, had moved to live in Israel. Andrew Cohen was there too, and Bob Cooper, so that affected my decision. But with Bernard it is a special relationship. I certainly respect his wisdom, his insights, and views although they are often different than mine; and I find that this is mutual. So we often disagree; we talk about how we always start our conversations from disagreements, and then within a few minutes of conversation we find ourselves in a different place and arriving at some consensus which is anchored in a renewed perspective, so we learn a lot from one another and certainly listen to one another. So basically we have very good discussions and conversations and by now it is not a question of agreeing, but of learning, updating one another. I am often surprised about this close relationship, because we have very different personalities, beliefs, and personal lives and come from very different worlds and function in different contexts. Bernard is practicing religion and resides in a community in Jerusalem; I live at the center of Tel Aviv and am very urban and Jerusalem and Tel Aviv are worlds apart, but these things are irrelevant in relation to our professional connections about “language” and language testing and language policy. I really cherish this relationship and am sure that my years in Israel would have been much different had Bernard not been living there. And an important point to mention is that in difficult times in life, he was there as a supporter and as a real “*mensch*.”

But this relationship is interesting also from a gendered point of view, because sometimes, younger women view it as subversive and unequal. Even when I say, “I learn from him, I respect his wisdom,” I hear women say, “What is it that a woman is always learning from a man?” So it is true that has been the norm—a younger woman learns from an older man, although I am sure that Bernard would also say that he respects me and my work; so I think perhaps it does defy the stereotype. I often think that my motivation to argue with him and disagree actually was instrumental in creating this relationship. I find that often women are not courageous enough and do not take the risk of disagreeing and expressing independent voices; this also happens because of the academic hierarchy, as in most cases the careers of women are so closely dependent on men’s support. My daughter is also in academia, but I can see how women of her generation construct themselves differently. They are very concerned about their own position and are aware of their rights as women; they voice their views, they are more confident and do not hesitate to demand what they think they deserve. You know, it is very interesting for me that you even asked the question about Bernard Spolsky, I wonder what motivated it.

A: Well, he seems to have been very fundamental for your thinking. And he’s made you think in ways that you might not have. You’re willing to listen to him.

E: I am not sure this is in fact so; even positioning your question as “he’s made you think in ways that you might not have,” assumes some kind of fatherly, mentor–protégé relationship and that is not very accurate. Yes, he is older, yes, he has been in the profession longer, yes, he probably wrote a letter for my promotion, but the work I have done comes from me; of course we were both influential on one another. He was influenced by me and I was influenced by him; we both made the other think about issues of assessment in different ways. But it was certainly not a one-way direction, and I wish we would distance ourselves from defining relationships as man–woman, older–younger, and senior–junior—we tend to simplify and categorize relationships in academia and overlook the mutuality of relationships. Mostly both of our views were shaped

in our ACROLT¹⁵ meetings, the Academic Committee for Research on Language Testing in Israel, which we created together along with Andrew Cohen in the '80s. We had lots of time for discussions in these conferences, unlike the LTRC, and we would always bring up issues about the unintended consequences of tests. David Nevo, my colleague, was there too, and it helped that he did not come from language so we could see it all from a larger perspective of measurement and program evaluation. We could examine issues of ethicality, consequences, and responsibility, as these were always on the table to be discussed. Since Israel is a small country and many people know each other, we could see these things very clearly. You see how the Minister of Education operates and how decisions are being made and how these affect teachers and students. People would visit these conferences from Europe and participate in the discussions on consequences and washback of these tests on their educational systems.

Perhaps this is the moment to express my appreciation to the language testing group and how as a profession, along with my criticism, it had served such an important role in shaping us. For me the language testing profession is a lively and caring community; it is like a close family, and often more than that; in the worst times of my life I was lucky to know these people. The important thing is how people in this field with which you spent 30 years arguing, disagreeing, discussing, loving, and critiquing serve as such an important role in one's life. We should not forget that within these contestations and arguments there is great value in spending a professional career with people who live in different places in the world but still keep in contact. In this very group where we have disagreements and different directions we also have many meaningful connections that are a treasure; in a way we can afford to critique *because* we feel close.

A: Would you agree with this statement? You have been courageous in being nonconventional.

E: ((laugh)) Courageous, maybe, if it means saying what I think; that is probably true. Sometimes I just say it and have no control of my words. They just come out when you have passion about a field and you focus more on the consequences and effect of the words. So, is this what you mean by being courageous? Then yes, I say what I think just because I strongly believe the words may have an impact; when I see injustices I must react, I feel I cannot be silent. This is probably what leads my life and it is grounded in growing up in a very political family with a strong awareness of justice, or what they thought was justice, because justice has many faces as well. So when someone gives a paper and it is offensive to the community, I cannot just sit there and not react. I also very often react to express what I think others would have liked to say but they are not as courageous as me, as not everyone is in a privileged position as I am, especially minority students, so I often think it is my duty to react.

A: You've explained in detail how you've gotten from your 1978 dissertation on the cloze and FSI to where you are now. What is the link between what you did in 1978 and your talk tomorrow on Linguistic Landscape?

E: I would say it is more a development that occurred without much explicit planning; in fact nothing in my life is driven by conscious feeling but by curiosity and passion. So my interest in language policy certainly emerged from language testing as I realized how language tests are being used as policy tools. So once I realized that, I got interested in other mechanisms that serve the regime, the system to promote ideological agendas of those who have strong ideological

¹⁵Academic Committee for Research on Language Testing: <http://info.smkb.ac.il/home/home.exe/11571>

beliefs about language and nationalism. In my 2006 book *Language Policy*¹⁶ I examined a number of policy tools I called “mechanisms,” which implicitly affect and create de facto policies. Linguistic landscape—language in public space—is clearly used as a tool that is instrumental in creating such policies and it is used by central systems to determine power. My interest in linguistic landscape is related to my interest in the environment, the ecology, beyond what people use and say to how they represent themselves in space, which is somewhat different than how they represent themselves in casual conversations, or in writing letters or texts. This is also related to my deep interest into the new literacies, the expansion of the notion of language and the use of “linguaging” to express the multiple ways we communicate, through signs, and images, and sounds, and cyberspace. I am obviously also attracted to multilingualism and to the contact of different languages which exist in the public space, far more than to speech. You can see signs that are store names, announcements, and instructions that represent different languages embedded in one another—something you will never see on tests, which are still so monolingual and rarely include pictures or images. So linguistic landscape provides legitimacy to expanded varieties of languages, with fewer boundaries, in more open ways; perhaps it is the total antithesis of the languages of tests. These phenomena are mostly taking place in the public space, an area that has been mostly ignored in the past and is still more free. We can see how it is fine to be multilingual in public spaces, but not in the speech; that would be called code-switching. So linguistic landscape breaks the monolingual ideology and different theories explain it. I wish our tests would reflect that too as well as the rich languages one can find in the public space. I think about my granddaughter, little Alma, where does she get so much textual information? Sure, from books, but also from her surroundings, from signs in the public space; this is how she is learning to read. She’s looking around, she’s seeing signs—the city itself is a textbook. People are talking a lot now about using linguistic landscape for educational purposes. I believe it is a powerful method for language learning.

A: I looked at your book *Linguistic Landscape*,¹⁷ and it was very interesting to see the kinds of phenomena the authors were writing about—it is applied linguistics, but the first thing I thought about was semiotics, but it’s more than that because you come from it from language. There’s a different starting point.

E: It begins with language, but you cannot just stay there; doing language alone is not enough, so we are expanding the definition of language. But students love it: You go outdoors, you document language in the ecology, you see that language exists everywhere, not just in textbooks, and you become more aware of the environment and you are active, you take pictures, you analyze, you conclude, you understand new phenomena, you critique. You also learn how to react, protest, and guard the ecology as it belongs to all. It is a free space and people should be able to use it as well, not only big corporations or governments that create monuments where they are trying to market ideologies. Or if they do, let’s learn how to read these ideologies and understand how they are trying to manipulate the ecology, so we become sophisticated consumers of the ecology and also active agents who want to protect and change it.

A: On a more personal note, what are your hobbies?

¹⁶Shohamy, E. (2006). *Language policy: Hidden agendas and new approaches*. New York: Routledge.

¹⁷Shohamy, E., & Gortner, D. (Eds.). (2008). *Linguistic landscape: Expanding the scenery*. New York: Routledge.

E: ((laugh)) I just work all the time! No time for hobbies, work *is* my hobby. But there are at least two things I like to do in particular—one, which is directly related to the work on linguistic landscape, is photography, taking pictures of people especially when they interact and use languages, and the other is cycling.

A: With your fold-up bike.

E: Yeah, with my two Bromptons, the blue or the yellow, depending on my mood; I don't know why we call things hobbies, you just do these in a more natural way; you ride your bike to work and like doing it, does it mean it is "a hobby"? But the thing I like most is what I am doing right now with you, talking with people. Whether it is on the plane or on park benches in Tel Aviv, or in Singapore or in tapas bars in Barcelona, I just love talking to people, and almost always about language experiences. So my current research is "experiencing" language policy. In most multilingual countries, every person has a language story, and these stories are just fascinating and deep and each one different than the other. I talk mostly to older people who share with me their language experiences especially of learning Hebrew, and the cost of having to switch languages in the middle of their lives—what did they lose, what did they gain, how did the language change affect their lives and functioning and of course how did they learn the language.

A. Nick told me you have a fascination with pens. Is that true?

E: Pens, computers, cameras, gadgets. Here are some pens.

A: Do you use all of them?

E: I use of all of them, but I go through periods. Now I'm into this one. Now I'm into green ink, then black. I think a lot of it is psychological—I am always worried that I won't be able to write, so having pens gives me the illusion and maybe the hope that it will also generate ideas; getting a new laptop is also useful for that purpose.

A: Elana, to wrap up this interview, can you tell me about language testing conferences you particularly like?

E: One conference I really liked was the Cambridge ALTE in 2008, because it addressed issues of social responsibility and 600 people attended, mostly teachers who discussed their experiences in classroom assessment as part of their teaching. One reason why some conferences are not so satisfying is that the testing organizations, with all due respect to Cambridge ESOL and ETS, are often not such a good influence. Can you see a conference in medicine where you have people who are developing drugs presenting papers? So I am always suspicious when testing organizations present papers, although I must say the testing organizations can often play an important role in encouraging and promoting important themes such as professionalizing the field, educating teachers about assessment, raising awareness about creating good tests and especially about social concerns or the consequences of language citizenship tests. So I am often torn about these issues as well and try to believe that these organizations are doing good.

A: Are there important moments, books, papers, or presentations in language testing that have stayed with you over the years?

E: I think Spolsky's *Measured Words*¹⁸ was very influential. I was very impressed by Samuel Messick's work; it was an inspiration for me. I did not meet him personally, but I read his things

¹⁸Spolsky, B. (1995). *Measured words: The development of objective language testing*. Oxford, UK: Oxford University Press.

long ago and then heard him talk at AERA. I respected his ideas about test consequences so much, because I'd been thinking about these topics for a long time so it was very meaningful to hear and read his papers. George Madaus, Pierre Bourdieu, and Michel Foucault have also been influential; they provided legitimacy for my thoughts. It is meaningful when you read that such prominent scholars and philosophers think about things that you had in mind but were hesitant to express. Claire Kramsch and Merrill Swain are inspirational models of independent women with very sophisticated and deep insights. My good friend Alastair Pennycook with his critical views about language and about everything in fact; and Tim McNamara, his thinking, his views, his writings, and his activism have been very inspiring and embedded in a most meaningful friendship. I am indebted to them and to Bernard Spolsky, whom I already discussed. Another person I do not know personally but whose writings have been influential for my thinking is Louis Moll; he is not in language testing but he introduced the notion of "funds of knowledge," the appreciation of the knowledge that most people have but is not appreciated in the educational world, especially not in testing. Testers, via the tests, "sell" the ideology of what people should know, rather than building on what they do know. And as to influence, the most important funds of knowledge have come from my own wonderful daughters, Daphna and Orlee, who taught me so much about so many things, testing and language and life and justice and caring; this interview cannot suffice in describing all I have learned and are learning from them. And little Alma, it is not about me asking her questions but it's about Alma telling me about what she knows, what she thinks, and she knows so much; she is constantly asking me deeper questions to learn more and more. So often "I" do not have the answers to her sophisticated questions as I flunk these tests . . . perhaps children should test adults and not the other way around, if it will force adults to learn new things that are relevant in the current world.

A: Do you have a book or publication that you are most proud of?

E: Probably the *Power of Tests*. I felt good about it because I was writing about a phenomenon that was such a central part of testing but that somehow was overlooked, so I felt I was doing something new then. I am especially happy that I had the opportunity to read the acknowledgement, which shows how my daughters contributed to the shaping of my ideas, to Orlee, a week before the terrible accident. I like *Language Policy* in a very different way because it was after Orlee died and it brings her into the book in many ways, in the introduction about "Languaging with Orlee," how we continue to create Orlee's life in her death through languaging in multiple forms. She is included in the main theme of the book where I quote lines from her diary in reference to language and life; these are two entities that cannot be controlled whereas there are always those who try to control them, via language policies and tests; but language and life have their own energy and dynamics and refuse to be controlled, not by policies, nor by tests as they have their own dynamics, pace and energy. So being able to share and connect the personal with the professional in this book and erasing the artificial boundaries of work and life, is a major agenda, as they are all one.

A: Elana, you were awarded the ILTA Lifetime Achievement Award recently. Your reaction?

E: I clearly appreciate the recognition, but would I like to end with a quote from one of my closest friends who wrote: . . . "I'm so glad for you that this has been acknowledged. It matters. And if they hope that by stressing the lifetime achievement bit (you're done all you can), they're hoping you may retire or be nice to them for a change, we know that's not going to happen."

ELANA SHOHAMY'S PUBLICATIONS IN LANGUAGE TESTING

Books

- Shohamy, E., & Walton, R. (Eds.). (1992). *Assessment for feedback: Testing and other strategies*. Dubuque: IA: Kendall/Hunt.
- Shohamy, E. (2001). *The power of tests: A critical perspective of the uses of language tests*. Singapore: Longman.
- Shohamy, E. (2006). *Language policy: Hidden agendas and new approaches*. London: Routledge.
- Shohamy, E., & Hornberger, N. H. (Eds.). (2008). *Encyclopedia of language and education, Volume 7: Language testing and assessment* (2nd ed.). New York: Springer.

Journal Articles

- Shohamy, E. (1982a). Affective considerations in language testing. *Modern Language Journal*, 66, 3–17.
- Shohamy, E. (1982b). The cloze procedure and its applicability for testing Hebrew as a foreign language. *Hebrew Annual Review*, 6, 101–114.
- Shohamy, E. (1982c). Predicting speaking proficiency from cloze tests: Theoretical and practical considerations for test substitutions. *Applied Linguistics*, 3, 161–171.
- Politzer, R., Shohamy, E., & McGroarty, M. (1983). Validation of linguistic and communicative language tests for Spanish-English bilingual programs. *Bilingual Review*, 10(1), 3–20.
- Shohamy, E. (1983a). Rater reliability of the oral interview speaking test. *Foreign Language Annals*, 16, 219–222.
- Shohamy, E. (1983b). The stability of the oral proficiency trait on the oral interview speaking tests. *Language Learning*, 33, 527–539.
- Shohamy, E. (1984). Does the testing method make a difference? The case of reading comprehension. *Language Testing*, 1, 147–170.
- Nevo, D., & Shohamy, E. (1986). Evaluation standards for the assessment of alternative testing methods: An application. *Studies in Educational Evaluation*, 12, 149–158.
- Shohamy, E., Reves, T., & Bejerano, Y. (1986). Introducing a new comprehensive test of oral proficiency. *English Language Teaching Journal*, 40, 212–220.
- Shohamy, E. (1988). A proposed framework for the testing of oral proficiency. *Studies in Second Language Acquisition*, 10, 165–179.
- Shohamy, E. (1989). Evaluating competence and performance in the foreign language classroom. *English Teachers' Journal*, 39(1), 20–26.
- Shohamy, E., Gordon, C., Kenyon, D., & Stansfield, C. (1989). The development and validation of a semi-direct test for assessing oral proficiency in Hebrew. *Bulletin of Higher Hebrew Education*, 4(1), 23–44.
- Shohamy, E. (1990). Language testing priorities: A different perspective. *Foreign Language Annals*, 23, 385–393.
- Shohamy, E. (1991a). Connecting testing and learning in the classroom and on the program level. In J. K. Phillips (ed.), *Building bridges and making connections. Northeast Conference Reports*, (pp. 154–178). So. Burlington, VT: Northeast Conference.
- Shohamy, E. (1991b). Discourse analysis and language testing. *Annual Review of Applied Linguistics*, 11, 115–131.
- Shohamy, E. (1991c). The Hebrew speaking test: An example of international cooperation in test development and validation. *AILA Review*, 7, 83–95.
- Shohamy, E. (1991d). International perspectives of language testing and systems and policy. *ACTFL Annual Review of Foreign Languages*, 16, 91–107.
- Shohamy, E., & Inbar, O. (1991). Construct validation of listening comprehension tests: The effect of text and question type. *Language Testing*, 8, 23–40.
- Shohamy, E. (1992). Beyond proficiency testing: A diagnostic feedback testing model for assessing foreign language learning. *Modern Language Journal*, 76, 513–521.
- Shohamy, E., Gordon, C., & Kraemer, R. (1992). The effect of raters' background and training on the reliability of direct writing tests. *Modern Language Journal*, 76, 7–33.
- Shohamy, E. (1994). The validity of direct versus semi-direct oral tests. *Language Testing*, 11, 99–123.
- Shohamy, E. (1995). Performance assessment in language testing. *Annual Review of Applied Linguistics*, 15, 188–211.

- Shohamy, E., Donitza-Schmidt, S., & Ferman, I. (1996). Test impact revisited: Washback effect over time. *Language Testing*, 13, 298–317.
- Shohamy, E. (1997). Testing methods, testing consequences: Are they ethical? Are they fair? *Language Testing*, 14, 340–349.
- Shohamy, E. (1998). Critical language testing and beyond. *Studies in Educational Evaluation*, 24, 331–345.
- Jennings, M., Fox J., Shohamy, E., & Graves, B. (1999). The effect of choice on test takers' scores in language tests. *Language Testing*, 15, 426–456.
- Shohamy, E. (2000). Interfaces of language testing and second language acquisition, revisited. *System*, 28, 541–553.
- Shohamy, E. (2001). Democratic assessment as an alternative. *Language Testing*, 18, 373–391.
- Shohamy, E. (2007). Language tests as language policy tools. *Assessment in Education*, 14, 117–130.
- Levin, T., & Shohamy, E. (2008). Achievement of immigrant students in mathematics and academic Hebrew in Israeli school: A large scale evaluation study. *Studies in Educational Evaluation*, 34, 1–14.
- McNamara, T., & Shohamy, E. (2008). Language tests and human rights. *International Journal of Applied Linguistics*, 18(1), 89–95.
- Shohamy, E. (2008). Overview: Language policy and language assessment: The relationship. *Current issues in Language Policy and Language Planning*, 9, 363–373.
- Shohamy, E., & Kanza, T. (2009). Language and citizenship in Israel. *Language Assessment Quarterly*, 6, 83–87.

Book Chapters

- Shohamy, E. (1981). Inter-rater reliability of the oral interview and concurrent validity with cloze procedure in Hebrew. In A. S. Palmer, P. J. M. Groot, & G. Trosper (Eds.), *The construct validation of communicative tests: An introduction* (pp. 94–105). Washington, DC: Teachers of English to Speakers of Other Languages.
- Shohamy, E. (1983a). Inter-rater and intra-rater reliability of oral interview and concurrent validity with cloze procedure. In J. Oller (Ed.), *Issues in language testing research* (pp. 229–236). Rowley, MA: Newbury House.
- Shohamy, E. (1983b). Six methods for testing reading comprehension. In S. Jones, M. Wesche, M. DeBrisay, & T. Paribakht (Eds.), *Proceedings of the 5th Language Testing Colloquium* (pp. 77–94). Ottawa, Canada: Carlton University.
- Shohamy, E. (1984). Input and output in language testing. In T. Culhane, C. Klein-Braley, & D. Stevenson (Eds.), *Occasional paper, practice and problems in language testing* (pp. 159–166). Colchester, UK: University of Essex.
- Shohamy, E. (1985). An alternative model for assessing language proficiency at the end of secondary school. In V. Kohonen, C. Klein-Braley, & H. von Essen (Eds.), *Practice and problems in language testing*, 8 (pp. 149–156). Tampere, Finland: Finnish Association for Applied Linguistics.
- Shohamy, E., Reves, T., & Bejerano, Y. (1985). From testing research to educational policy: Introducing a new comprehensive measure of oral proficiency. In V. Kohonen, H. von Essen, & C. Klein-Braley (Eds.), *Practice and problems in language testing*, 8 (pp. 133–144). Tampere, Finland: Finnish Association for Applied Linguistics.
- Shohamy, E. (1992). The connection between testing and learning. In E. Shohamy & R. Walton (Eds.), *Assessment for feedback: Testing and other strategies* (pp. 1–28). Dubuque, IA: Kendall/Hunt.
- Shohamy, E. (1993a). A collaborative/diagnostic feedback model for testing foreign languages. In D. Douglas & C. Chapelle (Eds.), *A new decade of language testing research* (pp. 185–202). Arlington, VA: Teachers of English to Speakers of Other Languages.
- Shohamy, E. (1993b). The exercise of power and control in the rhetorics of testing. *Carleton papers in applied language studies X* (pp. 48–63). Ottawa, Canada: Carlton University.
- Shohamy, E. (1993c). The exercise of power and control in the rhetorics of testing. In A. Huhta, K. Sajavaara, & S. Takala (Eds.), *Language testing: New openings* (pp. 23–38). Jyväskylä, Finland: University of Jyväskylä Institute for Educational Research.
- Shohamy, E. (1993d). The power of tests: The impact of language tests on teaching and learning. *NFLC Occasional Papers*. Washington, DC: NFLC.
- Shohamy, E. (1994a). Issues of language planning in Israel: Language and ideology. In R. Lambert (Ed.), *Language planning around the world: Contexts and systemic change* (pp. 131–142). Washington, DC: National Foreign Language Center.

- Shohamy, E. (1994b). The role of language tests in the construction and validation of second language acquisition theories. In E. Tarone, S. Gass, & A. Cohen (Eds.), *Research methodology in second language acquisition* (pp. 133–142). Hillsdale, NJ: Erlbaum.
- Shohamy, E. (1994c). The use of language tests for power and control. In J. Alatis (Ed.), *Georgetown University Round Table on language and linguistics* (pp. 57–72). Washington, DC: Georgetown University Press.
- Shohamy, E. (1995). Language testing: Matching assessment procedures with language knowledge. In M. Birenbaum & F. Dochy (Eds.), *Alternatives in assessment of achievements, learning processes and prior knowledge* (pp. 142–160). Boston: Kluwer.
- Shohamy, E. (1996). Competence and performance in language testing. In G. Brown, K. Masmliaer, & K. Williams (Eds.), *Performance and competence in second language acquisition* (pp. 136–151). Cambridge, MA: Cambridge University Press.
- Shohamy, E. (1998a). Alternative assessment in language testing: Applying a multiplism approach. In E. Li & G. James (Eds.), *Testing and evaluation in second language education* (pp. 99–114). Hong Kong: Lang.
- Shohamy, E. (1998b). Assessment in early age: Foreign language abilities of early language learners. In M. Met (Ed.), *Critical issues in early second language learning* (pp. 185–197). Glenview, IL: Scott Foresman-Addison-Wesley.
- Shohamy, E. (1998c). How can language testing and second language acquisition benefit from each other: The case of discourse. In L. Bachman & A. Cohen (Eds.), *Interfaces between second language acquisition and language testing research* (pp. 156–176). Cambridge, UK: Cambridge University Press.
- Shohamy, E. (1998d). Inside the “black box” of classroom language tests. *Studies Anglica Posnaniensia*, 33, 343–352.
- Shohamy, E. (1998e). Language tests as de facto educational and linguistic policies. In V. Berry & A. McNeill (Eds.), *Policy and practice in language education* (pp. 23–42). Hong Kong: The University of Hong Kong Press.
- Shohamy, E. (1999). Fairness in language testing. In A. Kunnan (Ed.), *Fairness and validation in language assessment* (pp. 85–89). Cambridge, UK: Cambridge University Press.
- Connor-Linton, J., & Shohamy, E. (2001). Register variation, oral proficiency sampling and the promise of multi-dimensional analysis. In S. Conrad & D. Biber (Eds.), *Variations in English: Multi-dimensional studies* (pp. 124–138). Essex, UK: Pearson Education Limited.
- Shohamy, E. (2001a). The relevance of language testing to second language acquisition. In H. Pürschel & U. Raatz (Eds.), *Tests and translation: Papers in memory of Christine Klein-Braley* (pp. 103–114). Bochum, Germany: AKS-Verlag.
- Shohamy, E. (2001b). The responsibility of the language tester. In R. Cooper, E. Shohamy, & J. Walter (Eds.), *New perspectives and issues in educational language policy. Papers in honor of Bernard Spolsky* (pp. 113–132). Amsterdam, the Netherlands: John Benjamins.
- Shohamy, E. (2004). Assessment in multicultural societies: Applying democratic principles and practices to language testing. In B. Norton & K. Toohey (Eds.), *Critical pedagogies and language learning* (pp. 72–93). Cambridge, UK: Cambridge University Press.
- Shohamy, E. (2005). The power of tests over teachers: The power of teachers over tests. In D. Tedick (Ed.), *Second language teacher education: International perspectives* (pp. 101–110). Mahwah, NJ: Erlbaum.
- Shohamy, E. (2006a). The power of language tests, the power of the English language and the role of ELT. In J. Cummins & C. Davison (Eds.), *International handbook of English language teaching* (pp. 521–532). New York: Springer.
- Shohamy, E. (2006b). Re-thinking assessment for advanced language proficiency. In H. Byrnes, H. Weger-Guntharp, & K. A. Sprang (Eds.), *Educating for advanced foreign language capacities: Constructs, curriculum, instruction, assessment* (pp. 188–208). Washington, DC: Georgetown University Press.
- Levin, T., & Shohamy, E. (2007). The role of academic language in understanding the mathematics achievements of immigrant students in Israel. In C. S. Sunal & K. Mutua (Eds.), *The enterprise of education: Research on education in Africa, the Caribbean, and the Middle East* (pp. 313–336). Tuscaloosa, AL: Information Age.
- Levin, T., Shohamy, E., & Inbar, O. (2007). Achievements in academic Hebrew among immigrant students in Israel. In E. Olshtain & N. Nevo (Eds.), *Hebrew in a global world*. Israel: Magnes.
- Shohamy, E. (2007). Tests as power tools: Looking backwards, looking forward. In J. Fox et al. (Eds.), *Language testing reconsidered* (pp. 141–152). Ottawa, Canada: University of Ottawa Press.
- Inbar-Lourie, O., & Shohamy, E. (2009). Assessing young language learners: What is the construct? In M. Nikolov (Ed.), *Contextualizing the age factor: Issues in early foreign language learning* (pp. 83–96). Berlin: Mouton de Gruyter.
- Shohamy, E. (2009). Language tests for immigrants: Why language? Why tests? Why citizenship? In G. Hogan-Brun, C. Mar-Molinero, & P. Stevenson (Eds.), *Discourses on language and integration* (pp. 45–59). Amsterdam: John Benjamins.
- Shohamy, E., & Kanza, T. (2009). Citizenship, language and nationality in Israel. In G. Extra, M. Spotti, & P. Van Avermaet (Eds.), *Language testing, migration and citizenship: Cross-national perspectives* (pp. 242–260). London: Continuum.