Identifying second language speech tasks and ability levels for successful nurse oral interaction

with patients in a linguistic minority setting: An instrument development project

Talia Isaacs [a,d]

Michel D. Laurier [b]

Carolyn E. Turner [a,d]

&

Norman Segalowitz [c,d]

[a]McGill University, Montréal [b]Université de Montréal, Montreal
[c]Concordia University, Montréal,   [d]Centre for the Study of Learning and Performance

Identifying second language speech tasks and ability levels for successful nurse oral interaction with patients in a linguistic minority setting: An instrument development project

Abstract

One of the most demanding situations for members of linguistic minorities is a conversation between a health professional and a patient, a situation that frequently arises for linguistic minority groups in North America, Europe and elsewhere. The present study reports on the construction of an oral interaction scale for nurses serving linguistic minorities in their second language (L2). A mixed methods approach was used to identify and validate a set of speech activities relating to nurse interactions with patients and to derive the L2 ability required to carry out those tasks. The research included an extensive literature review, the development of an initial list of speech tasks, and validation of this list with a nurse focus group. The retained speech tasks were then developed into a questionnaire and administered to 133 Quebec nurses who assessed each speech task for difficulty in an L2 context. Results were submitted to Rasch analysis and calibrated with reference to the Canadian Language Benchmarks, and the constructs underlying the speech tasks were identified through exploratory and confirmatory factor analyses. Results showed that speech tasks dealing with emotional aspects of caregiving and conveying health-specific information were reported as being the most demanding in terms of L2 ability, and the most strongly associated with L2 ability required for nurse-patient interactions. Implications are discussed with respect to the development and use of assessment instruments to facilitate L2 workplace training for healthcare professionals.

(232 words)

Identifying second language speech tasks and ability levels for successful nurse oral interaction

with patients in a linguistic minority setting: An instrument development project

There is growing interest in the nature of health communication (Candlin & Candlin, 2003) and, in particular, communication problems that can arise when patients and caregivers do not speak the same first language (L1) (Robinson, 2002). One concern about such language gaps is that communication may suffer to the point of compromising healthcare quality. For example, a language barrier may make it difficult to establish good caregiver-patient rapport and, as a consequence, the patient may receive less attentive care and experience less satisfaction with the medical treatment than patients for whom there is no language barrier (Carrasquillo, Orav, Brennon, & Burstin, 1999; Jacobs, Chen, Karliner, Agger-Gupta, & Mutha, 2006). In some areas of healthcare—including the delivery of bad news (e.g., Gillotti, Thompson & McNeilis, 2002), obtaining informed consent (e.g., Meyer, 2003; Schenker, Wang, Selig, Ng, & Fernandez, 2007), or talking about death, suicide, pain or other sensitive issues—good communication with the patient can be especially critical for obvious emotional, ethical and legal reasons (Anderson et al., 2003; Chen, Youdelman, & Brooks, 2007). More generally, there can be communication problems between patient and health professionals even with the same L1 (e.g., Cegala et al., 2008; Ong, de Haes, Hoos, & Lammes, 1995), and these problems are likely exacerbated when one speaker has a poor command of the language used.

In linguistically diverse countries such as the United States and Canada and especially in those with several *official* language communities (e.g., Canada, India), providing equitable healthcare services in multiple languages can be a challenge (Bélanger, 2003; Bowen, 2001; Canadian Medical Association, 2005; Lavizzo-Mourey, 2007). Often this entails health professionals speaking in their non-fluent second language (L2) or enlisting the aid of an interpreter (frequently a medically unqualified family member or hospital personnel; Fernandez & Schenker, 2010; Lee et al., 2006). Such situations can also result in the patient accommodating linguistically to the caregiver rather than the other way around. It can be

expected, therefore, that increasingly there will be occasions when caregivers do not feel they have the necessary L2 competence to handle their patients' communicative needs with the sensitivity and comprehensiveness they deserve.

Steps can be taken to address such language barriers (Schenker, Lo, Ettinger, & Fernandez, 2008). One is to identify those situations that caregivers perceive to be difficult to handle in an L2 and to target the relevant sociolinguistic and pragmatic functions in L2 training. While more is becoming known about the nature of communication between patients and caregivers with the same L1 (e.g., Roter & McNeilis, 2003; see also papers in Heritage & Maynard, 2006a), relatively little is known about such communication when the health professional uses a non-fluent L2. Systematic research on this topic is sorely needed. Having a tool capable of identifying areas of communication that caregivers view as presenting significant barriers in using their L2 would enable researchers to focus on the most vulnerable aspects of medical communication. This should aid the development of interventions to overcome language barriers. The present study investigated a way to identify areas of communication where such language barriers are likely to exist between nurses and patients.

Before going into the study, however, it will be useful to consider how language barriers might manifest themselves. One approach to this is provided by speech act theory (Austen, 1962; Searle, 1969, 1976). The challenge nurses face when using the L2 is *not* simply to translate L1 words and expressions into their L2. Rather, they must communicate intentions (perform illocutionary rather than just locutionary or utterance acts). People carry out illocutionary acts in a wide variety of ways, including by using complex forms of indirect speech (e.g., "Do you think you might be able to do something about that smoking habit?"). Formal speech act analysis requires careful consideration of the *context* in communication takes place. Moreover, as Blum-Kulka (1997) and Cooren (2005) have shown, such formal analysis also requires taking into account the *sequencing* of the conversational exchange. Thus, analyses cannot be restricted to isolated speech samples examined in the abstract. Because illocutionary

acts are expressed in conventionalized ways that vary across linguistic cultures (Blum-Kulka, 1997), language barriers likely arise at this level. It is important, therefore, to understand how speech acts are normally expressed in the target language and to identify the kinds of miscommunication and misunderstanding that arise from failure to handle speech acts appropriately in the L2.

Conversational analysis (Drew, 2005; Heritage & Maynard, 2006b) provides another perspective on language barriers. Here the focus is the "negotiation" that characterizes communication. If the caregiver has low L2 proficiency or is unfamiliar with L2 conversational routines and conventions, then the natural flow of the conversational negotiation may be compromised, creating a language barrier. Yet another approach to language barriers comes from an intercultural pragmatics perspective (Wierzbicka, 2008). Here, culturally specific semantic scripts are held to underlie the meanings of certain words (e.g., for key English words like "personal" and "privacy"). Failure to appreciate the semantic script underlying key words can result in language barriers.

Thus, for any broadly defined speech activity (e.g., conveying bad news about a prognosis, or obtaining informed consent) there are required speech acts, rules for conversational negotiation, and cultural scripts that must be handled properly in the L2 to avoid linguistic barriers. Such intercultural differences in language use have implications for L2 training (Knutson, 2006) aimed at overcoming barriers. An important first step, however, is to systematically identify the communicative activities in which language barriers are likely to arise. Once the potential locus of a language barrier has been found, the underlying nature of the barrier can be investigated, the results of which should inform remedial action.

To summarize, the goal of overcoming language barriers in healthcare settings minimally requires three steps—identifying communicative tasks where significant language barriers are likely to arise; closely examining those tasks using various analytical tools; and developing appropriate remedial steps. The present study focuses on the first step. Rather than targeting communicative tasks *a priori* using a

particular theoretical framework (e.g., Burnard, 2003), a qualitative (inductive) approach was used to start the process, with the results subsequently submitted to quantitative analyses. For the qualitative phase, domain experts (i.e., nurses) were consulted about their language use to guide the selection of communicative activities to be examined. The literature on nurse-patient communication has been growing (Shattell, 2003) and includes case studies documenting what nurses actually say in particular situations (Skott & Eriksson, 2005; Wadensten, 2005). No studies that we are aware of, however, have systematically investigated nurses' speech activities using an L2, nor their perceptions of task relevance and difficulty. The present research aims to fill this void. In a second phase, the data were analyzed quantitatively (Rasch analysis; factor analyses) to expand upon and refine the qualitative results. The main emphasis of this study is on this second, quantitative phase.

Research design and methodology

The motivation for this project arose from within the healthcare community and the desire of health practitioners to linguistically accommodate their patients. Direct communication with patients as opposed to reliance on third parties is also in the researchers' interest. The general target group is in-service registered nurses who work in their L2 when serving linguistic minorities in Quebec, Canada.

The goal of the study is threefold: (1) to identify and validate a set of speech tasks that nurses regard as the most relevant when serving patients in their non-dominant language, (2) to identify the dimensions underlying these speech tasks, in order to better understand the L2 skills needed to carry them out, and (3) to calibrate the nursing-specific speech tasks onto a previously-validated rating scale commonly used for workplace language training, in order to extend the applicability of the scale to nursing-specific communicative tasks. This study is conceived of as a stepping stone to the eventual development of an oral assessment instrument designed to enhance L2 teaching and learning of health-specific oral content for nurses. The subject of this paper is the development and validation of a list of

relevant speech tasks, the first step in this longer-term research agenda. The specific research questions are:

RQ1: To what extent do nurses identify, in a consistent and reliable way, specific speech tasks that they perceive as relevant and requiring an appropriate level of language ability in L2 interactions with patients?

RQ2: To what extent can the speech tasks so identified be scaled in relation to existing L2 performance benchmarks?

RQ3: To what extent do the patterns of speech tasks identified by nurses fall into well-defined categories that could serve as the basis for developing a pedagogically oriented assessment instrument for nurse L2 training?

The study used a sequential exploratory mixed methods design commonly used for instrument development (Creswell & Plano Clark, 2007). The qualitative and quantitative data were collected sequentially across two stages. Decisions made at the first stage (qualitative) influenced the direction of the second stage (quantitative). In addition, the first stage results helped guide the second stage in instances of quantitative analysis limitations. The qualitative methods included a literature review, focus group, and verbal protocol. The product of this first stage was a questionnaire for nurses. The quantitative methods included Rasch analysis and factor analyses of the questionnaire data. The product of the second stage was a list of speech tasks and ability levels that are specific and relevant to the nursing profession.

Methodology and results for the qualitative stage (QUAL)

As part of the questionnaire development, we wanted to align the speech tasks with a well-known, previously validated oral interaction scale commonly used to assess adult L2 speakers. The *Canadian Language Benchmarks* (*CLB*), a descriptive L2 proficiency scale available in English and French, was

chosen for this purpose (Pawlikowska-Smith, 2000). Proficiency levels are indicated as reference points (benchmarks) on a 12-point scale grouped into three categories: Beginner, Intermediate and Advanced.

Preliminary questionnaire

The qualitative stage began with an extensive literature review that centered on identifying nursing-specific speech tasks. The goal was to generate a list of speech tasks for a preliminary questionnaire to be used with a focus group of nurses. Twenty-two speech tasks were identified and a preliminary questionnaire was developed with scales for *relevance* and *difficulty*.

Five L1 French nurses representative of diverse care domains in a Quebec hospital ($M_{experience} = 27.8$ years; 20–35) participated in a 2 hour focus group led by two of the researchers. Using the preliminary questionnaire as a guide, the principal goals were to identify relevant speech tasks from the list provided; gather specific contextual examples; estimate the linguistic difficulty necessary to perform these speech tasks in their L2 (i.e., English); and add to the list (e.g., nurses deemed the speech task "rephrasing patients' ideas" to be relevant and "rather difficult" to perform in an L2 due to its lexical demands). This process generated 19 speech tasks. Some were deleted due to irrelevancy, and new relevant ones were added.

Verbal protocol

In order to refine the list of speech tasks to be included in the questionnaire, a verbal protocol was administered to an experienced L1 French nurse. This was pursued as an alternative to piloting for validation purposes. Using a think aloud procedure, the nurse was audio recorded verbalizing her thoughts in French as she reflected on the language level needed to perform each speech task in her L2 (English) using the *CLB* scale. At the end of the process, she reflected, "if the task is of a technical nature, then telling them [patients] what they have to do is simple. But uh... it's much more complex when you have to deal with the patient's emotions." The data were transcribed and subjected to a

qualitative content analysis. The 19 speech tasks from the focus group were retained with slight wording modification. The English version of the speech tasks appears in Appendix A.

Methodology and results for the quantitative stage: QUAN

Up to this point (i.e., the first stage), we had identified relevant speech activities for nurse-patient interaction. Next, we developed a final questionnaire. The goal was to reach out to a larger population of nurses to validate relevancy and identify the proficiency level needed to perform each speech task. In considering the *CLB* descriptors and content analysis of the verbal protocol, it was evident that the lower and higher levels included language functions less relevant to nursing. We decided to align our proficiency levels for the final questionnaire with the middle seven levels; thus the *CLB* levels 4 to 10 became our levels 1 to 7. Two formats of a final questionnaire were developed to accommodate participants: a hard copy bilingual questionnaire, and an online questionnaire with two versions (French/English) developed through the online questionnaire development tool, *Survey Monkey* (http://www.surveymonkey.com). Participants were provided with the 19 speech tasks and the 7-level *CLB* descriptors and were asked to identify the relevant *CLB* level for each speech task using the prompt, "The person who can perform the following task…should at least be at level."

The participants who completed the questionnaire were 159 registered nurses (reduced to 133 after further analyses, see below), who were familiar with the day-to-day realities of the healthcare context and were, therefore, deemed well-positioned to make judgments about language use in their profession ($M_{experience}$ = 19.2 years; 1–35). They were recruited from language schools, the Quebec nurse licensing body, hospitals, a conference and a local university. The highest sampling came from the first two groups (79 nurses from nine different language schools and 30 from the licensing body). In addition, variation by geographical region was achieved, with nurse respondents from 11 out of the 17 regions in Quebec. From self-reported data we learned of their language use at work. Since the majority were L1 French speakers (n = 110), these data will be reported here. On a scale of 1 to 4 (1 = *never*; 4 = *always*),

the mean French use at work was 3.89 (*SD* = .31); the mean English use was 2.22 (*SD* = .58); and the mean use of another language was 1.20 (*SD* = .40). Finally, their mean L2 English proficiency, as reported on a 3-point scale (*1 = beginner, 3 = advanced*) was 1.71 (*SD* = .66).

Descriptive statistics were calculated using SPSS 14.0. Rasch analysis was then employed using FACETS 3.60.0 to refine the data, determine the relevance of the 19 speech tasks, and align them with the *CLB* scale for L2 proficiency level. Finally, factor analyses were performed using SAS 9.1, in order to identify the underlying constructs being measured.

Indicator Description and Modeling

The questionnaire data were analyzed for the purpose of determining the degree of agreement among the nursing community about the 19 speech tasks. From this perspective, a good indicator (i.e., speech task) needed to show convergence among the nurses. This convergence was observed for each indicator in reference to the following criteria:

- Correspondence between the level previously identified by the members of the focus group and the nurse questionnaire respondents' estimates of the necessary level to perform a speech task;

- Differences between respondents in their individual estimates: when the nurses agree, they tend to assign the same level to a given speech task;

- Consistency of their ratings in relation with other indicators: the speech task ranking in terms of difficulty is similar from one respondent to the other.

In order to identify the indicators that show convergence, we used a combination of descriptive statistics (mean, *SD*, correlation) and Rasch analysis, a methodology used in previous studies (e.g., Blais, Laurier & Rousseau, 2009). Using Rasch analysis software, we applied a polychotomous Rasch model to analyze the ordered categories obtained from the rating scale (Andrich, 1999). Seven categories corresponding to the seven proficiency levels were recognized using a simple 2-facet model (Indicators x Respondents). This type of analysis allows one to model the data instead of just describing

it. The difficulty indices are more robust than the mean scores because they are less sample dependent. In addition, indices about the fit of the indicators and respondents' (nurses') ratings in relation to the model can yield valuable information about the consistency of their response patterns.

Rasch analysis is iterative, and three iterations were needed to clean up the data. This process resulted in 26 nurse respondents being removed from the original 159, due either to invariant ratings across speech tasks or erratic (unmodeled) rating behaviour. The final sample size, therefore, was 133 nurses. Generally, one would expect a significant proportion of the items (indicators) to be rejected due to poor model fit. Interestingly, all 19 speech tasks were validated. That is, nurse respondents did agree on what should and should not be considered difficult tasks. We mainly considered infit mean squares due to their utility for identifying indicators that may threaten the validity of the measurement system and ease of interpretation. This index is based on the chi-square statistic with each observation weighted by its model variance. The expected value is 1. Values below 1 indicate overfit which, in this situation, is unproblematic, as some redundancy is expected. Values above 1 indicate underfit because of noise or sources of variance that add on the main variance, reflecting variations in relation with the trait being measured. When the fit index is above 2, we can conclude that the indicator or the respondent may corrupt the measurement system. The analysis revealed that all indicators had infit indices below Wright and Linacre's (1994) suggested cut-off value of 1.4. Thus, all items were retained.

The process of calibration aims at locating each facet element on the "Rasch ruler." This common scale usually ranges from – 4 to + 4 logits. The analysis is normally set to fix the 0 point at the mean item difficulty (in this case, the speech task), and speech tasks and raters (in this case, nurse respondents) are located accordingly. Because the Rasch model uses logits as the measurement unit, the scale is considered arithmetic (e.g., an indicator calibrated at 2 is twice as difficult as an indicator calibrated at 1). The placement of respondents on the ruler is interpreted as the degree of severity of their ratings.

In Figure 1, the 1st column (column R) represents the "Rasch ruler." Only the − 2 to + 2 range is shown because few observations fall outside these limits. The column L on the right shows where the boundaries between levels are located. The most difficult speech tasks as judged by the nurses appear at the top of the "items" column, while the least difficult appear at the bottom. As predicted, the indicator, "Inform patient of bad news" is clearly the most difficult task to perform by a nurse in an L2, whereas "Check identity over phone" was the easiest overall. The other speech tasks are grouped within the −1 and +1 range. The far right column evidences the wide distribution of the nurses (raters) in relation to severity.

<PLEASE INSERT FIGURE 1 ABOUT HERE>

Ideally an indicator should be neither too difficult nor too easy. For a given speech task, all questionnaire respondents should identify the same difficulty level. In practice, this situation never occurs because raters never reach perfect consensus. When the respondents tend to agree on a level, the distribution around the mean is not skewed, the mode is obvious and the standard deviation is low. Figure 2 shows such a distribution for the indicator,"Making encouraging statements to a patient" The mean (3.79), the Rasch measure (0.89) and the mode (4) confirm the nurses' proposal to consider that the level required to perform this speech task is Level 4 (Level 7 of the *CLB*). The low standard deviation (1.49) can be interpreted as general agreement among the nursing community, and this hypothesis is sustained by the infit mean square, which is close to 1.

<PLEASE INSERT FIGURE 2 ABOUT HERE>

In summary, the Rasch analysis confirmed that all speech tasks were functioning adequately from a statistical standpoint (i.e., without undue distortion) and should be retained in the subsequent analyses. At the same time, the nurses whose fit indices were not predicted by the model—that is, whose responses were either too erratic or unvaried to be modeled—were removed from the data set. This served to "clean up" the data so that patterns could be more easily discerned in the subsequent analyses.

Factor Analyses

The questionnaire data were subjected to an exploratory factor analysis using the varimax rotation. Multiple indices suggested that three factors should be extracted. First, the first eigenvalue was preeminent, accounting for 73.5% of the variance; however, the specks in the scree plot did not level off until the fourth eigenvalue. This suggests that the first three eigenvalues accounted for a substantial proportion of the variance (93.6% cumulatively). Second, the application of Kaiser's criterion, which holds that only eigenvalues greater than one should be retained, led to the selection of three eigenvalues (Stevens, 2002). Finally, the overall root mean square residual was less than .05, which implies that the reproduced matrix with three factors represented the original data well.

Following the selection of three factors preliminarily, the data were subjected to a confirmatory factor analysis. Items with high residuals are usually discarded in this procedure in the interests of attaining better model fit and for reasons of parsimony. However, because the Rasch item indices and descriptive statistics had provided evidence that all 19 speech tasks were functioning adequately for the purpose of developing a nursing-specific descriptive scale, it was decided *a priori* that no speech tasks would be discarded in the factor analyses. Furthermore, the focus group nurses had deemed these speech tasks to be the most the relevant nurse-patient interactive tasks. This substantive argument overrode any statistical expedience of excluding speech tasks. Several models were assessed under the constraint of retaining all speech tasks, and the second order confirmatory factor analysis model shown in Figure 3 was found to be the best. One higher order factor (F4) predicted three second order factors (F1, F2, and F3), all of which, in turn, predicted the 19 speech tasks.

<PLEASE INSERT FIGURE 3 ABOUT HERE>

A series of goodness-of-fit statistics were examined to assess the validity of the model. Schermelleh-Engel, Moosbrugger, and Müller's "rule of thumb" ranges of good model fit were adopted as a guideline (2003). In fact, none of the indices approached the good fit range. The *p* value for the χ^2

test, for example, was < .0001 (good fit range: $.05 \leq p \leq .10$), the *RMSEA* estimate was .11 (good fit range: $0 \leq RMSEA \leq .05$), *CFI* was.89 (good fit range: $.97 \leq CFI \leq 1.0$), and *NNFI* was .85 (good fit range: $.97 \leq NNFI \leq 1.0$). These fit indices would clearly not be acceptable for high-stakes test development (e.g., a nurse licensing exam). However, the result was deemed sufficient for the overall goal of deriving a descriptive rating instrument designated for pedagogical use, provided that the speech task clustering was readily interpretable in the process of labeling the factors.[1]

The speech tasks with the highest factor loadings (shown in brackets) for factor one (F1) were: managing a patient's anger or impatience (.94); informing a patient of bad news (.81); refusing unreasonable requests made by a patient (.80); making apologies to a patient (.79); showing empathy towards a patient (.79); and reassuring a patient (.70). The descriptions "anger," "apologies," "empathy," "reassuring," and managing a patient's likely adverse reaction to receiving bad news and potentially negative reaction to having an unreasonable request refused suggests a factor that deals with managing a patient's emotions. F1 was, therefore, labeled, "*Emotional aspects of caregiving*."

The speech tasks with the highest factor loadings for factor two (F2) were: giving instructions to a patient accurately on the phone (1.02) or face to face (.99); rephrasing/confirming a patient's description of pain (.80); summarizing a patient's health situation (.80); and providing information about available services in relation to a patient's medical condition (.75). Giving instructions or information and summarizing, paraphrasing, or acknowledging a patient's description of pain seemed to delineate a factor that entails conveying information about a patient's medical condition. F2 was, therefore, labeled, "*Factual aspects related to the patient's medical condition*."

The speech tasks with the highest factor loadings for factor three (F3) were: checking a patient's identity over the phone (1.02); giving instructions on how to get to the hospital over the phone (.94); refusing unreasonable requests made by a patient (.85); and making encouraging statements to a patient (.84). "Checking a patient's identity," giving directions over the phone, and offering encouragement

seemed to relate to communicating with a patient on non health-specific topics. Notably, the speech task "refusing unreasonable requests" also loaded onto F1, the emotional factor. This double loading can perhaps be explained by the scenario where refusing a patient's unreasonable request could potentially escalate into needing to deal with his/her emotions. Conversely, refusing a patient's request to smoke on the premises, the example that was provided with the speech task (see Appendix A), might be considered quite routine or uninvolved. F3 was, therefore, labeled, "*Routine aspects*."

Finally, the factor loadings between F4 and the second order factors were as follows. The loading onto F1, *Emotional aspects of caregiving*, was .76. The loading onto F2, *Factual aspects related to the patient's medical condition*, was .55. Finally, the loading onto F3, *Routine aspects*, was –.23. Both F1 and F2 were significantly associated with F4 at the $p < .01$ level (Stevens, 2002). Thus, factual aspects and emotionally charged aspects of caregiving, both of which are health-specific, relate to the overarching construct. F3, however, was not significantly correlated with F4. In fact, the ability represented by F3 appears to run contrary to the development of F4, in that the correlation was weak and negative. This may be because the formulaic language involved in carrying out speech tasks such as "checking a patient's identity" or "making encouraging statements" has little to do with the L2 ability required to respond to emotionally-sensitive health topics or convey specific information about a patient's health. F4 was, therefore, labeled, "*Health-specific verbal interaction with patients*."

Correspondence with CLB Bands

Having completed the Rasch and factor analyses, what remained was to align the 19 speech tasks with the *CLB* scale bands so that these nurse-patient interactive tasks could be directly associated with a *CLB* level. First, each speech task was linked with the second order factor with the highest factor loading (F1, F2, or, F3). Next, Rasch item indices and descriptive statistics were examined to assign each speech task to the appropriate *CLB* level. Although the reduced *CLB* scale with seven levels was used in the questionnaire administered to nurses, Figure 4 shows that the indicators were only calibrated

onto five *CLB* scale bands. That is, the resulting scale was a 5-point scale. Notably, the easiest items were all associated with F3, the routine factor. Due to their generic (i.e., non health-specific) nature, these speech tasks may not be instructive for getting at health-specific nurse-patient communication, the construct of interest (F4). The most salient speech tasks in this regard (i.e., those most strongly associated with F1 or F2), were aligned with *CLB* scale bands 7–9.

<center><PLEASE INSERT FIGURE 4 ABOUT HERE></center>

Discussion

Summary and Contributions

The purpose of this study was to develop an oral interaction scale for nurses serving linguistic minorities in their L2. Qualitative findings informed the development and implementation of a quantitative instrument in a sequential manner. In the qualitative stage, an extensive literature review on oral interactions between health practitioners and patients and input from focus group nurses led to a draft list of the most pertinent speech activities. This list was developed into a bilingual questionnaire, piloted through verbal protocol, and distributed to nurses working in different contexts across Quebec. In the quantitative stage, the questionnaire data were analyzed through descriptive statistics and Rasch modeling and adequate item functioning was confirmed. Next, the factor analyses, with all 19 speech tasks included as exogenous variables, resulted in grouping the indicators into three dimensions which, in turn, were underpinned by the broader dimension of health-specific oral interaction skills for nurses (F4). Speech tasks dealing with a patient's emotions (F1) and conveying health-specific information (F2) were the most pertinent. The speech tasks related to routine aspects of nurses' professional life (F3), on the other hand, bore little relation to the construct of interest.

That emotionally-charged speech tasks are generally more difficult to carry out in an L2 than routine tasks to carry out in an L2 may not surprise either domain experts or "lay" people outside the healthcare system. Indeed, people often make intuitive judgments about the relative difficulty of the

tasks they carry out in their daily lives and extrapolate from experience. A main contribution of the study is that it provides an *empirical* basis for determining which speech tasks are the most difficult and relevant. While nurses are not linguists, they were arguably the best placed to gauge the emotional impact of an interactive task and how that might render a particular speech task easier or more difficult to carry out in an L2, a perspective that should be taken into account in the L2 training and testing of nurses. The perspectives of other health professionals (e.g., doctors, pharmacists) could similarly be probed using a mixed methodology.

Notably, the speech tasks were calibrated onto the *CLB* band descriptors. This national standard has exerted an influence on Canadian classrooms that train adult L2 learners for professional purposes, including newcomers seeking to integrate into the healthcare sector. Aligning the speech tasks with the *CLB* bands extends the applicability of the scale to healthcare professionals. In addition, it provides information on the *CLB* level necessary for nurses to be able to carry out interactive tasks in their professional domain (*CLB* levels 7–9 for the health-specific speech tasks in this study). This complements other attempts to devise assessment tools for Canadian nurses using the *CLB* like the *CELBAN*, which was created for licensure purposes (CELBAN, 2010).

The descriptive scale developed in this study is part of a longer-term research agenda to devise an assessment tool to help health practitioners improve their L2 skills when dealing with minority language patients in the workplace. Conceived of as a tool for teaching and learning, this instrument will include assessment tasks that center on emotional aspects of caregiving and factual aspects related to patients' medical conditions. The challenge will be to design an instrument that can be easily integrated into typical classroom and self-training activities. The speech tasks could be expanded into a series of role plays, for example, which have been a mainstay for the training of healthcare professionals (e.g., Walker, Cedergren, Trofimovich, Gatbonton, & Mikhail, 2008; Watt & Lake, 2007). The skill

development resulting from such an instrument could enhance nurses' willingness to engage in using their L2 in situations they might otherwise shy away from and, thus, better serve minority communities.

Limitations

The interplay between qualitative and quantitative methods is an advantage of the present study. However, several limitations need to be taken into account in interpreting the results. First, the focus group was relatively small; a more extensive focus group consultation may have yielded a larger set of candidate communicative activities to examine. However, Creswell and Plano-Clark (2007) emphasize that in sequential exploratory mixed-methods designs, a small sample size is appropriate in the initial qualitative exploration. Second, the test sample was limited to in-service nurses from different regions of Quebec. Results are, to some extent, reflective of the Quebec healthcare needs and system. The study would need to be replicated to determine how well the findings generalize. Obtaining data from a larger sample of nurses would also have been desirable. This is particularly the case for the factor analyses, since the study fell short of the norm of 10 participants for each factor. However, the use of qualitative data arguably compensated for this shortcoming. Third, the study was limited to nurses' perspectives of their language needs. A necessary next step is to conduct a needs analysis to gauge the perspectives of other health professionals charged with different daily communicative tasks (e.g., doctors), with the long-term goal of enhancing workplace-specific L2 training. The "Essential Skills" inventory developed by Human Resources and Skills Development Canada may be a useful starting point in this regard (de Vries, 2009). Finally, the study investigated nurses' language needs when serving patients in their L2 with a focus on English or French, Canada's official languages. The choice of these languages understates the multilingual reality of Quebec society (Lamarre, 2003), however, and ignores contexts in which health practitioner training in non-official languages may be beneficial (e.g., Inuit and First Nations languages). Future research could focus on instrument development to enhance nurse training in additional languages.

Concluding Remarks

This study reported a mixed methods approach for systematically identifying healthcare-relevant speech tasks and placing them in an ordered list. The ranking of these nurse-patient interactive tasks reflects both nurses' perceptions of the linguistic difficulty of the task, and nurses' self-assessments of their ability to effectively communicate in their L2 in that task situation. The resulting list of speech tasks can inform fine-grained follow-up analyses of linguistic barriers aimed at supporting the development of L2 training for nurses. In such follow-up analyses, it must be remembered that skill in using the L2 is not just a matter of knowing how to translate from L1 to L2. As Heritage and Maynard (2006b) point out, "it is by acting together that doctor [caregiver] and patient assemble each particular visit with its interactional textures, perceived features and outcomes. … [The] ordinary norms and practices of language use and social interaction exert a powerful and systematic influence on the texture and features of medical visits, and do so in fine detail" (pp. 19-20). As pointed out earlier, how a communicative event unfolds when the caregiver is speaking in the L2 is interesting from a variety of linguistic standpoints (speech act theory, conversational analysis, intercultural pragmatics, etc.). Understanding how caregivers perceive the challenges inherent in participating in these events is a crucial first step in the process of developing appropriate L2 training for caregivers.

Footnotes

[1] As outlined above, letting the numbers drive the data by "throwing out" speech tasks was not pursued in this study. This was due to the nature of the low-stakes instrument being developed and to the importance placed on the domain experts' input in the development and validation of the list of speech tasks. However, in the interests of examining whether the data *could* support a better fitting model if statistically redundant speech tasks were to be removed, models that optimized fit by excluding speech tasks with large unique variance components were examined. A three factor solution with nine speech tasks yielded the best fitting model, and all indices fell within Schermelleh-Engel et al.'s "good fit" range ($p = .22$; $RMSEA = .04$; $CFI = .99$; $NNFI = .99$). This suggests that an assessment instrument for nurse licensure purposes could potentially be created using the same data in light of fit statistics; however, this was not the purpose of the present study.

Acknowledgments

References

Anderson, L., Scrimshaw, S., Fullilove, M., Fielding, J., Norman, J. (2003). Culturally competent

healthcare systems: A systematic review. *American Journal of Preventative Medicine, 24,* 68–79.

Andrich, D. (1999). Rating scale analysis. In J. P. Keeves and G. N. Masters (Eds.), *Advances in*

measurement in educational research and assessment (pp. 110–121). Oxford: Pergamon.

Austin, J. L. (1962). *How to do things with words.* Cambridge, MA: Harvard University Press.

Bélanger, M. (2003). *L'accès aux soins de santé pour les communautés minoritaires de langue officielle*.

Ottawa: Chambres des Communes, Canada.

Blais, J. -G., Laurier, M. D., & Rousseau, C. (2009). Deriving proficiency scales from performance

indicators using the Rasch model. In E. V. Smith Jr., & G. E. Stone (Eds.), *Criterion referenced*

testing: Practice analysis to score reporting using Rasch measurement models (pp. 528–540).

Maple Grove, MN: JAM Press.

Blum-Kulka, S. (1997). Discourse pragmatics. In T. A. van Dijk (Ed.), *Discourse as social interaction,*

Volume 2, (pp. 38–63). London: SAGE Publications.

Bowen, S. (2001). *Language barriers in access to health care*. Ottawa: Health Canada.

Burnard, P. (2003). Ordinary chat and therapeutic conversation: Phatic communication and mental

health nursing. *Journal of Psychiatric and Mental Health Nursing, 10*, 678–682.

Canadian Medical Association. (2005). *National physician survey*: *New data reflect multiculturalism's*

impact on medicine. Retrieved January 18, 2010, from

http://cma.ca/index.cfm/ci_id/ 10025394/la_id/1.htm

Candlin, C., & Candlin, S. (2003). Health care communication: A problematic site for applied linguistics

research. *Annual Review of Applied Linguistics, 23,* 134–154.

Carrasquillo, O., Orav, J., Brennon, T., & Burstin, H. (1999). Impact of language barriers on patient

satisfaction in an emergency department. *Journal of General Internal Medicine, 14,* 82–87.

Cegala, D. J., Bahnson, R. R., Clinton, S. K., David, P., Gong, M. C., Monk III, J. P., et al. (2008). Information seeking and satisfaction with physician–patient communication among prostate cancer survivors. *Health Communication*, *23*, 62–69.

CELBAN. (2010). *Canadian English Language Benchmark Assessment for Nurses*. Retrieved January 18, 2010, from http://www.celban.org/celban/display_page.asp?page_id=1

Chen, A. H., Youdelman, M. K., & Brooks, J. (2007). The legal framework for language access in healthcare settings: Title VI and beyond. *Journal of General Internal Medicine*, *22*, 362–367.

Cooren, F. (2005). The contribution of speech acct theory to the analysis of conversation: Hoe pre-sequences work. In K. L. Fitch & R. E. Sanders (Eds.), *Handbook of language and social interaction,* (pp. 21–40). Mahwah, NJ: Lawrence Erlbaum.

Creswell, J. W., & Plano Clark, V. L. (2007). *Designing and conducting mixed methods research*. Thousand Oaks, CA: Sage.

de Vries, E. (2009). Workplace essential skills in practice: A Canadian perspective. In R. Maclean & D. Wilson (Eds.), *International handbook of education for the changing world of work* (pp. 2713–2730). The Netherlands: Springer.

Drew, P. (2005). Conversational analysis. In K. L. Fitch & R. E. Sanders (Eds.), *Handbook of language and social interaction,* (pp. 71–102). Mahwah, NJ: Lawrence Erlbaum.

Fernandez, A., & Schenker, Y. (2010). Time to establish national standards and certification for health care interpreters. *Patient Education and Counseling 78*, 139–140

Gillotti, C., Thompson, T., & McNeilis, K. (2002). Communicative competence in the delivery of bad news. *Social Science & Medicine*, *54*, 1011–1023.

Heritage, J., & Maynard, D. W. (Eds.). (2006a). *Communication in medial care: Interaction between primary care physicians and patients.* Cambridge: Cambridge University Press.

Heritage, J., & Maynard, D. W. (2006b). Introduction: Analysing interaction between doctors and

    patients in primary care encounters. In J. Heritage & D. W. Maynard, (Eds.), *Communication in*

    medial care: Interaction between primary care physicians and patients (pp. 1–21). Cambridge:

    Cambridge University Press.

Jacobs, E., Chen, A., Karliner, L. S., Agger-Gupta, N., & Mutha, S. (2006). The need for more research

    on language barriers in health care: A proposed research agenda. *The Milbank Quarterly*, *84,* 111–

    133.

Knutson, E. M. (2006). Cross-cultural awareness for second/foreign language learners. *Canadian*

    Modern Language Review, 62, 591–610.

Lamarre, P. (2003). Growing up trilingual in Montreal: Perceptions of college students. In R. Bayley &

    S. R. Schecter (Eds.), *Language socialization in bilingual and multilingual societies* (pp. 62–80).

    Clevedon, UK: Multilingual Matters.

Lavizzo-Mourey, R. (2007). Improving quality of US health care hinges on improving language

    services. *Journal of General Internal Medicine*, *22*, 279–28*0.*

Lee, K. C., Winickoff, J. P., Kim, M. K., Campbell, E. G., Betancourt, J.R., Park, E. R., et al. (2006).

    Resident physicians' use of professional and nonprofessional interpreters: A national survey.

    Journal of the American Medical Association, *296*, 1050–1053.

Meyer, B. (2003). Bilingual risk communication. *Arbeiten zur Mehrsprachigkeit, 50.* Universität

    Hamburg.

Ong, L., de Haes, J., Hoos, A., & Lammes, F. (1995). Doctor-patient communication: A review of the

    literature. *Social Science and Medicine, 40,* 903–918.

Pawlikowska-Smith, G. (2000). *Canadian Language Benchmarks 2000*. Ottawa, ON: Centre for

    Canadian Language Benchmarks.

Robinson, M. (2002). *Communication and health in a multi-ethnic society.* Bristol: Policy Press.

Roter, D., & McNeilis, K.S. (2003). The nature of the therapeutic relationship and the assessment and consequences of its discourse in routine medical visits. In T. Thompson, A. Dorsey, K. Miller, & R. Parrott (Eds.), *Handbook of health communication* (pp. 121–140). Mahwah, NJ: Lawrence Erlbaum.

Schenker, Y., Lo, B., Ettinger, J. D., & Fernandez, A. (2008). Navigating language barriers under difficult circumstances. *Annals of Internal Medicine, 149,* 264–269.

Schenker, Y., Wang, F., Selig, S. J., Ng, R., & Fernandez, A. (2007). The impact of language barriers on documentation of informed consent at a hospital with on-site interpreter services. *Journal of General Internal Medicine*, 22, 294–299.

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, 8, 23–74.

Searle, J. (1969). *Speech acts.* Cambridge, UK: Cambridge University Press.

Searle, J. (1976). A classification of illocutionary acts. *Language in Society, 5,* 1–23.

Shattell, M. (2003). Nurse-patient interaction: A review of the literature. *Journal of Clinical Nursing*, 13, 714–722.

Skott, C., & Eriksson, A. (2005). Clinical caring: The diary of a nurse. *Journal of Clinical Nursing*, 14, 916–921.

Stevens, J. P. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). Mahwah, NJ: Lawrence Erlbaum.

Wadensten, B. (2005). The content of morning time conversations between nursing home staff and residents. *Journal of Clinical Nursing*, 14, 84–89.

Walker, N., Cedergren, H. J., Trofimovich, P., Gatbonton, E., & Mikhail, E. (2008). *Someone to talk to: A virtual patient for medical history interview training in a second language*. INTED (International Technology, Education, and Development Conference) Proceedings, Valencia, Spain.

Watt, D., & Lake, D. (2007). *Medical communication assessment project (M-CAP): An Alberta project for international medical graduates*. Report prepared for the M-CAP, Faculty of Medicine, University of Calgary.

Wierzbicka, A. (2008). A conceptual basis for intercultural pragmatics and world-wide understanding. In M. Pütz & J. Neff-van Aertselaer (Eds.), *Developing contrastive pragmatics* (pp. 3–45). Berlin: Mouton de Gruyter.

Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, *8*, 370. Retrieved May 19, 2010, from http://www.rasch.org/rmt/rmt83b.htm

Appendix A

List of Speech Tasks Identified as Most Relevant by Quebec Nurses

1. Checking a patient's identity over the phone.

2. Giving instructions on how to get to the hospital or clinic over the phone.

3. Rephrasing/confirming a patient's description of pain.

4. Giving instructions to a patient accurately face to face (e.g., regarding medication).

5. Giving instructions to a patient accurately on the phone (e.g., regarding medication).

6. Providing additional information about available services in relation to the patient's

   condition (e.g., after surgery).

7. Summarizing a patient's health situation to the patient.

8. Summarizing/rephrasing a patient's feelings in reaction to a diagnosis.

9. Describing to a patient in detail the common professional interventions (e.g., blood tests,

   injections).

10. Reasoning with a patient (e.g., after surgery has been delayed).

11. Reassuring a patient (e.g., in an emergency situation, before surgery).

12. Showing empathy towards a patient.

13. Clarifying to a patient what the doctor tried to explain regarding his/her medical condition.

14. Asking routine questions to a patient.

15. Refusing unreasonable requests made by a patient (e.g., can't smoke).

16. Making apologies to a patient (e.g., right before an intervention that might be

   painful/uncomfortable, forgot to do something requested by patient).

17. Making encouraging statements to a patient (e.g., when patient beginning to walk again).

18. Managing a patient's anger or impatience.

19. Informing a patient of bad news (e.g., condition has become worse).

R	Indicators		L	Respondents
+2				..
				....
				..
				......
	19) Inform patient of bad news			.
			6	...
				...
				.....
----	-----------------------------------	-----------------------------------	----	.....
+1				.....
	13) Clarify doctor's message			....
	18) Manage patient's anger	8) Describe reaction to diagnosis		.....
	10) Reason with patient			.....
	7) Summarize health situation		5	....
	11) Reassure patient			....
	6) Give info about services			...
----	-----------------------------------	-----------------------------------	----	....
	12) Show empathy	5) Give instructions on phone		....
0				....
	4) Give instructions in person			....
	9) Describe interventions		4	..
	16) Apologize to patient			.....
	15) Refuse unreasonable requests	3) Reformulate pain description		.....
----	-----------------------------------	-----------------------------------	----	.....
				...
	17) Encourage patient			.
	14) Ask routine questions		3	
-1				.
	2) Give directions over phone			
----	-----------------------------------	-----------------------------------	----	.
				.
			2	.
	1) Check identity over phone			
-2	-----------------------------------	-----------------------------------	---	

Figure 1. Location of the Indicators on the "Rasch Ruler"

Note. Only non-extreme scalar responses (scale levels 2-6) are shown here.
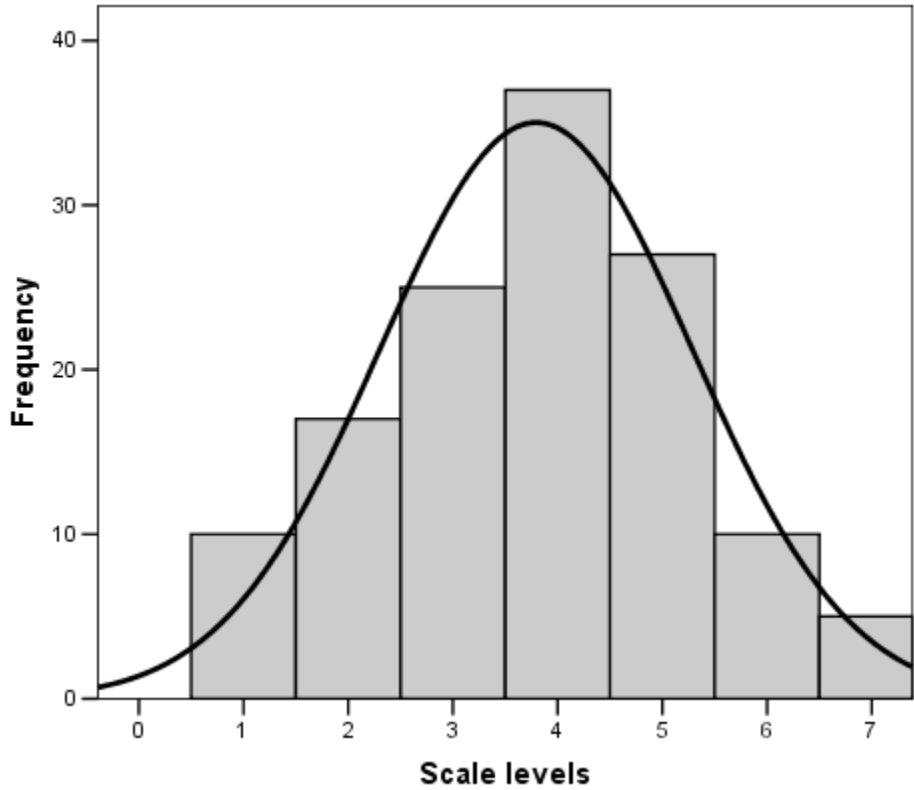
Item 17: ***Making encouraging statements to a patient (e.g., patient beginning to walk again).***

Rasch Item Indices
Difficulty: –0.89
Infit (MnSq): 1.07

Sample Statistics:
n = 131 (out of 133)
M = 3.79
SD = 1.49
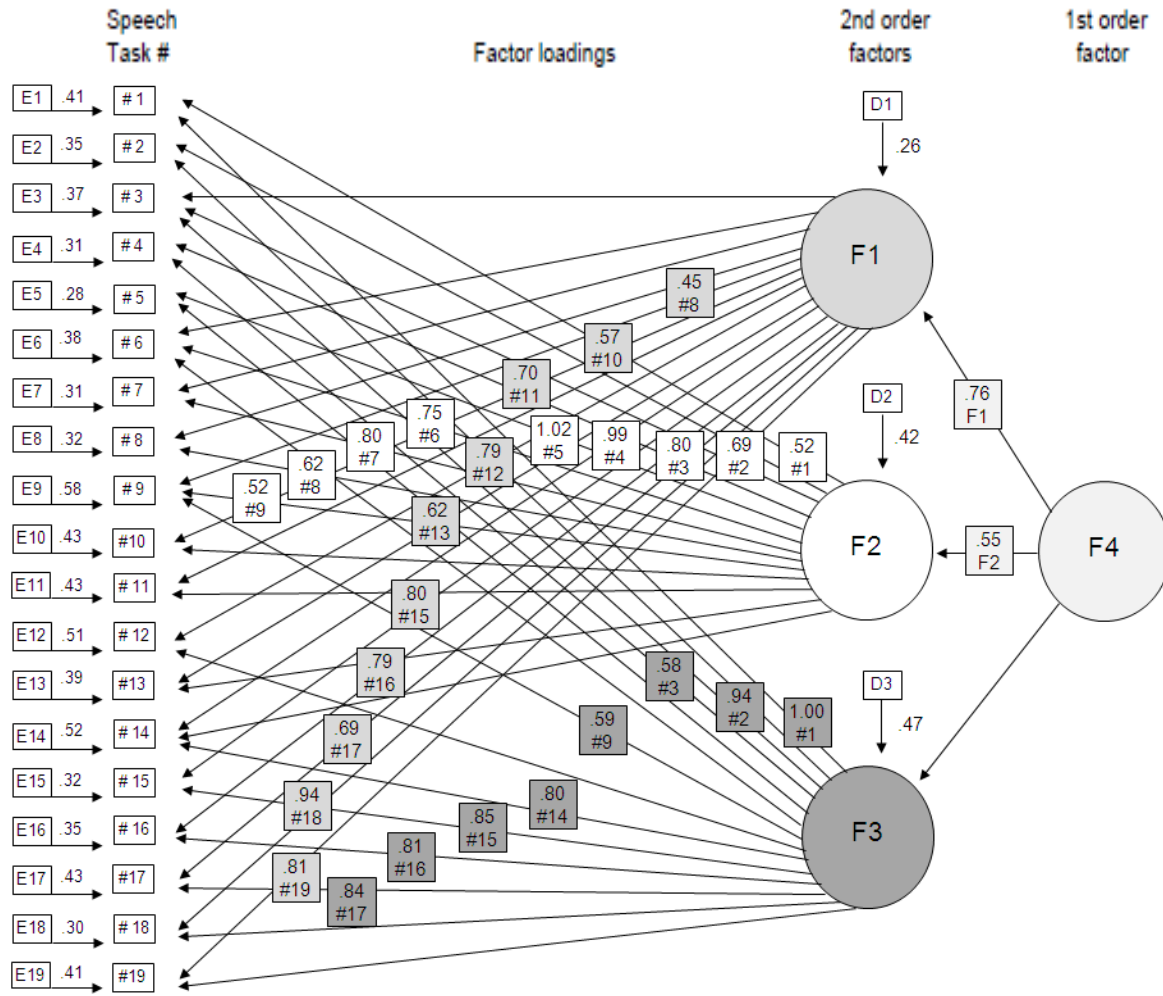
Figure 2. Indicator with an Adequate Distribution

Figure 3. Second Order Confirmatory Factor Analysis Model

Note. While the arrows show all of the relationships specified in the structural equations for the confirmatory factor analysis model, only statistically significant factor loadings are shown in the factor loading boxes (α = .01, two-tailed). The number appearing in the second line of these boxes after the '#' sign refers to the speech task item number that the factor is linked with. Boxes that contain 'D's are the error terms associated with the second order factors; boxes that contain 'E's are the error terms associated with the speech tasks. Notably, there is no correlation between the second order factors (i.e., they are orthogonal).
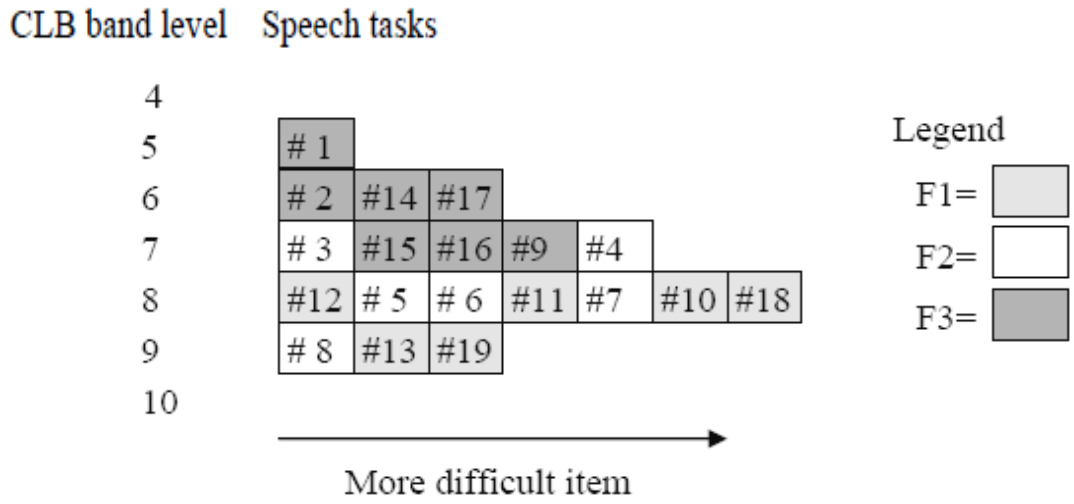
CLB band level   Speech tasks

Figure 4. Aligning speech tasks with CLB Scale Bands