A B S T R A C T

The frequent use of non-representative, time-consuming and low hit rate sampling methods in the study of minority cultures is of concern. Given the increasing cultural diversity of the Canadian population and the growing Chinese population, a sampling method using phonological rules to distinguish Chinese surnames from non-Chinese ones was developed.

Chinese surnames, transcribed according to their pronunciations, follow specific phonological rules. A flowchart was developed using the inclusion and exclusion criteria derived from these rules. Methods to validate Chinese surnames (identified from telephone directory) were external reference source, expert panel, and telephone interview.

A total of 266 possible Chinese surnames were selected using the flowchart. External reference source validated 153 surnames, expert panel 53, and telephone interview 23. An overall hit rate of 86.1% (229/266) was achieved, reflecting the validity of the phonological criteria in identifying individuals with a Chinese surname from the general population.

A B R É G É

L'utilisation fréquente de techniques d'échantillonnage non représentatives, fastidieuses et à faible taux de succès pour l'étude des cultures minoritaires est préoccupante. Vu la diversité culturelle grandissante de la population canadienne et la croissance de la population chinoise, nous avons élaboré une technique d'échantillonnage à base de règles phonologiques pour distinguer les noms de famille chinois des autres noms de famille.

Les noms chinois, transcrits selon leur prononciation, obéissent à des règles phonologiques précises. À partir de critères d'inclusion et d'exclusion dérivés de ces règles, nous avons mis au point un ordiogramme, puis validé des noms de famille chinois (trouvés dans l'annuaire du téléphone) au moyen d'un ouvrage de référence externe, d'un groupe d'experts et d'une entrevue téléphonique.

En tout, 266 noms pouvant être d'origine chinoise ont été sélectionnés à l'aide de l'ordiogramme. L'ouvrage de référence externe a permis d'en valider 153, le groupe d'experts, 53, et l'entrevue téléphonique, 23. Notre taux de succès global a été de 86,1 % (229/266), ce qui confirme la validité des critères phonologiques pour l'identification des personnes portant un nom de famille chinois dans la population.

How to Find Chinese Research Participants: Use of a Phonologically Based Surname Search Method

Erin Yuet Tjam, PhD, MHS, SLP (reg)

Research focussed on cultural minorities is timely because of the rapid demographic changes and increasing cultural diversity of the Canadian population. This diversity contributes to biological, psychological, social, functional and health variations that are important considerations for many fields of research.^{1,2} On the other hand, the frequent use of non-representative samples (e.g., culturally distinct neighbourhoods, minority culture organization records, snowballing), time-consuming and low hit rate sampling methods (e.g., random digit dialling, census survey information) in search of participants from minority cultures is of concern.³⁻⁷ Non-representative sampling affects generalizability, hence the application of research findings is limited. Sampling methods that are time consuming and have low hit rates can be very costly and ineffective, which creates financial and practical barriers to the prompt completion of a study.

The increasing heterogeneity of Canada's population is largely a result of the changing immigration patterns over the last century. Prior to the 1960s, immigrants were primarily European. By the late 1980s, substantial numbers came from South Asia, Southeast Asia, and China.⁸ The Asian population (especially Chinese) is projected to be the fastest-growing group among cultural minorities, and Chinese will constitute 23% of visible minorities.⁹ Latimer and Lundy reported that in the top 20 sources of immigration to Canada,

Hong Kong ranked first, with China, Vietnam, Korea and Taiwan following closely.¹⁰ These data suggest that the Chinese are the most common visible minority, and effective sampling strategies should be developed to accurately identify these individuals.

Using culturally distinct neighbourhoods like "China Town" or "Little Italy" allows researchers to conveniently identify individuals of interest from a culturally concentrated location.^{3,6} On the other hand, this method neglects a large group of individuals from the same culture who do not reside in these cultural neighbourhoods, and they may present a very different social, educational and financial background. Similarly, minority culture organizations (e.g., church, cultural centres) attract a sub-group of people whose interest and social background might be different from those not involved in the organization. Snowballing is a method whereby participating individuals suggest friends and neighbours who are of the same cultural background to be included in the sample, and they in turn suggest more participants for the sampling pool. This method faces the same restrictive and non-representative sampling limitations as the previous two and is likely to increase the homogeneity of the group respondents.

In an effort to improve representativeness, many studies used random digit dialling and census survey information.^{4,5} These methods improve representativeness because of the census coverage of the total population. However, trying to identify a small sub-cultural group from a large pool of the mainstream population is very time consuming and effortful. The number of times one correctly identifies the target group (i.e., hit rate) can be very low despite the large population base.

St. Mary's General Hospital, Kitchener, ON; Department of Health Studies and Gerontology, University of Waterloo, Waterloo, ON

Correspondence: Dr. Erin Y. Tjam, Director of Research, St. Mary's General Hospital, 911 Queen's Boulevard, Kitchener, ON, N2M 1B2

Financial support of this project provided through the Seniors Independence Program (SIP# 4687-06-93-458) and through a Canada's Drug Strategy - Community Health Researcher Award to Erin Y. Tjam (Grant #6606-5560-603).

In this study, a theoretically based, low cost, high hit rate, and representative approach of utilizing phonological methods to distinguish Chinese surnames from non-Chinese ones was developed to resolve the sampling shortcomings of other studies.

Theoretical background

Surname and ethnicity are strongly linked. Ethnicity refers to people who share the same ancestry or history, with distinctive patterns of family life, values, social norms, and *language*.¹¹ Surnames of people, which are language-based, can be used at the sampling stage as a selection criterion to identify individuals by race and name. This would identify an ethnic background rather than citizenship.

Chinese words are character-based and strokes are the basic units. To represent Chinese surnames (which are real words, mostly nouns and adjectives) in an alphabetical language, the method is to transcribe the surname according to its pronunciation. Phonological rules of Chinese language dictate that all Chinese words are monosyllabic.¹² Chinese surnames are therefore also monosyllabic (except for a few well-known two-word compound surnames). There are distinct phonological rules and patterns within each Chinese dialect that are quite different from most alphabetical languages.

In comparison with English phonetics, Chinese surnames – transcribed based primarily on the pronunciation of the Cantonese or Mandarin/Putonghua dialects (the two most common Chinese dialects) – can only begin with a vowel *sound* (e.g., /a, e, i, o/), or one of the following consonant *sounds*: /b, p, m, f, d, t, n, l, g, k, h, z, c, s, j, q, x, zh, ch, sh, y, w, r, ng, tz/. Further, Chinese surnames can only end with a vowel *sound* (e.g., /a, e, i, o, u, ou, ow, y, eu/), or one of the following consonant *sounds*: /m, n, ng, p, t, k/.^{12,13} These rules provide inclusion criteria for the phonologically based surname search method as only names that fall within the above conditions can be included (Table I). It is important to recognize that a *sound* refers to how a word is *pronounced*, not *spelled* (e.g., the name “Chieh” ends with a vowel *sound* “eh” even though its spelling is a consonant “h”,

TABLE I
Inclusion Criteria for Chinese Surnames based on the Cantonese and Mandarin/ Putonghua Phonological Rules

Phonological Structure	Inclusion Criteria
Number of Syllables	One Syllable
Initial Sounds	Any Vowels Consonant Sounds /b, p, m, f, d, t, n, l, g, k, h, z, c, s, j, q, x, zh, ch, sh, y, w, r, ng, hs, hw, tz, ts, sz/
Final Sounds	Any Vowels Consonant Sounds /m, n, ng, p, t, k/

TABLE II
Exclusion Criteria for Chinese Surnames based on the Cantonese and Mandarin/ Putonghua Phonological Rules

Phonological Structure	Exclusion Criteria
Number of Syllables	Multisyllabic
Initial Sounds	Consonant Cluster Sounds /bl, br, pl, pr, dr, dw, tr, gl, gr, kl, kr, sl, sp, st, sw, str, cl, cr, fl, fr, wr/ Fricative or Affricate Sounds /th, v/
Final Sounds	Fricative, Affricate or Liquid Sounds /s, f, v, sh, ch, dz, th, l, r/

therefore, “Chieh” fits the “vowel sound ending” condition).

Given these phonological patterns, any words/surnames that are multisyllabic (e.g., Hirdes, Curtis, Lubben); begin with a consonant cluster sound (e.g., Stones, Green); begin with a sound other than the given consonants (e.g., Thatcher, Yenn); or end with a fricative, affricate or liquid sound (e.g., Forbes, Rantz, Rose, Poole), will not qualify for possible Chinese words or surnames, and therefore can be excluded (see exclusion criteria in Table II).

A simple three-step flowchart (Figure 1) using the inclusion and exclusion criteria derived from the phonological rules can be used to decide whether a particular surname is Chinese. The same flowchart can be applied to first and middle Chinese names (which are also real words) for additional verification. Users can choose to follow either the flowchart's inclusion or exclusion path to reach a decision. This method enables researchers to take any population-based records (e.g., telephone directory, health records, census data) and efficiently distinguish Chinese surnames from the general population. As there are only a few hundred Chinese surnames,¹⁴ it is relatively easy to identify them.

METHODS

Sampling

The target population was telephone owners with a Chinese surname residing in the Waterloo region. Names of all telephone owners in the Waterloo region were obtained using a CD-ROM telephone directory. This method provided a near census population of all possible surnames (with the exception of unlisted telephone owners or individuals without a telephone, who represent a small percentage of the population). The two universities (University of Waterloo, Wilfrid Laurier University) in Waterloo house a large number of students from China, Taiwan and Hong Kong, hence the list of Chinese surnames is believed to be relatively comprehensive.

Data collection

Using the CD-ROM telephone directory, possible Chinese surnames were selected using the flowchart. To illustrate the mechanism of this process, the names “Lubben, Stone, Moss, Lian” were tested, either using the inclusion or exclusion path, and the decision rationale is summarized in Table

Figure 1. Flowchart for Identifying Chinese Surnames Using the Inclusion and Exclusion Path Based on the Cantonese and Mandarin/Putonghua Phonological Rules

III. Only “Lian” passed all three levels and reached the final level of being considered a Chinese surname. Surnames that passed all levels were listed in Table IV as possible Chinese surnames. They were then verified to see if they are true Chinese surnames.

Validation
Three methods were used to validate the phonologically based flowchart. The first method compared the list of possible Chinese surnames in Table IV against a validated external reference source, which

contained a list of Chinese surnames previously verified.¹⁵ Those surnames that corresponded with the validated reference source were categorized as true Chinese surnames. The second method used an expert panel. A group of 8 Chinese indi-

TABLE III Mechanism for Identifying Chinese Surnames Using the Flowchart and the Inclusion and Exclusion Criteria				
Flowchart Levels	Decision Process (using each criteria path) “Surnames”			
	“Lubben” (exclusion criteria)	“Stone” (inclusion criteria)	“Moss” (exclusion criteria)	“Lian” (inclusion criteria)
1st level (number of syllables)	<i>bi-syllabic name</i> - excluded at 1st level • cannot proceed	<i>monosyllabic name</i> - pass 1st level ➤ proceed to 2nd level	<i>monosyllabic name</i> - not excluded at 1st level ➤ proceed to 2nd level	<i>monosyllabic name</i> - pass 1st level ➤ proceed to 2nd level
2nd level (initial sound)		<i>consonant cluster /st/</i> - initial sound not a vowel or one of the sounds in Table I • cannot proceed	<i>bilabial initial sound /m/</i> - initial sound not a consonant cluster, fricative or affricate sound in Table II - not excluded at 2nd level ➤ proceed to 3rd level	<i>liquid initial sound /l/</i> - initial sound is a consonant in Table I - pass 2nd level ➤ proceed to 3rd level
3rd level (final sound)			<i>fricative final sound /s/</i> - final sound is a fricative - excluded at 3rd level • cannot proceed	<i>nasal final sound /n/</i> - final sound is a consonant in Table I ➤ proceed to decision level
Chinese Surname?	No	No	No	Yes

viduals who speak different dialects independently examined the surnames that were not on the external reference source to determine if they were Chinese surnames. When discrepancies occurred, open discussions followed and panel members presented their selection rationale. Their list was reviewed and the process was repeated until a consensus list was generated. Finally, individuals with surnames that could not be verified with the above two methods were telephoned directly. A simple question "is your last name a Chinese last name" was used to verify whether the surname was Chinese. This specific question was chosen to avoid confusion between ethnicity and nationality. The combined result of these three validation methods produced a list of true Chinese surnames.

Analysis

Descriptive analysis was employed to determine the percentage hit rate of the phonological method. The number of true Chinese surnames (n) was divided by the number of possible Chinese surnames (N) to generate a hit rate (Table V). Using phonological rules to identify Chinese research participants represents a theory-generating research, and the focus is on developing a theoretically sound method to efficiently distinguish Chinese surnames from non-Chinese surnames. Percentage hit rate is therefore a sufficient analytical method for this type of study.

RESULTS

A total of 266 possible Chinese surnames was selected using the flowchart based on the inclusion and exclusion criteria. Among the selected Chinese surnames, 153 were verified as true Chinese surnames using the external reference source. The expert panel agreed with the 153 surnames, and verified 53 more, yielding a total of 206 true Chinese surnames. Direct telephoning of the 60 unverified surnames yielded 23 additional true Chinese surnames. The total number of verified surnames selected from the phonological method was 229 of 266 (86.1% hit rate).

TABLE IV
A List of Chinese Surnames Identified In the Waterloo Region Using the Inclusion and Exclusion Criteria of the Phonological Method

Au	Fay	Jin	Liou	Qu	Wing
Bean	Feng	Joie	Liu	Sa	Won
Bui	Fok	Joo	Lo	Seng	Wong
Cai	Fong	Jung	Loi	Shen	Woo
Cao	Fu	Kam	Long	Shi	Wu
Cha	Fung	Kan	Loo	Shue	Xia
Chai	Funk	Kang	Lou	Sieu	Xiao
Chak	Gao	Kao	Low	Sim	Xie
Chan	Ge	Kat	Loy	Sin	Xing
Chang	Geng	Kau	Lu	Siu	Xiong
Chao	Gong	Kew	Lui	So	Xu
Chau	Gow	Kim	Luk	Sok	Yam
Che	Gu	King	Lum	Song	Yan
Chen	Guan	Kip	Lung	Soo	Yang
Cheng	Guo	Kit	Luo	Sop	Yap
Chiang	Guy	Ko	Luong	Su	Yau
Chieh	Ha	Kok	Ly	Sun	Ye
Chien	Hak	Kong	Ma	Sze	Yeo
Chim	Han	Koo	Mah	Ta	Yeung
Chin	Hang	Koy	Mai	Tai	Yim
Ching	Hau	Kuan	Mak	Tam	Yin
Chiu	He	Kun	Man	Tan	Yip
Cho	Heng	Kwan	Mau	Tang	Yiu
Choi	Heu	Kwok	Maw	Tao	Yo
Chong	Heung	Kwong	May	Teng	Yong
Choo	Ho	Ky	Mei	Teoh	Yoon
Choong	Hon	Lai	Mo	Tieo	Yu
Chow	Hong	Laing	Moh	Ting	Yuan
Chown	Hou	Lam	Moi	Tiong	Yuen
Choy	Hsieh	Lan	Mok	To	Yung
Chu	Hsiung	Lau	Monk	Tom	Zai
Chuang	Hsu	Law	Moon	Tong	Zee
Chui	Hsueh	Lee	Mua	Tsang	Zhai
Chun	Hu	Lem	Mui	Tse	Zhan
Chung	Hua	Leu	Nam	Tseng	Zhang
Cook	Huan	Leung	Neu	Tso	Zhao
Cui	Huang	Lew	Ning	Tsoi	Zheng
Dai	Hui	Li	Noi	Tsui	Zhong
Dong	Hung	Lian	Pak	Tu	Zhou
Du	Hwang	Liang	Pan	Tung	Zhu
Duong	Ing	Liao	Pang	Wan	Zuk
Egan	Ip	Lieu	Pau	Wang	
Ek	Jan	Lim	Peng	Way	
Eng	Jain	Lin	Ping	Wei	
Fan	Jiang	Ling	Pong	Wey	

TABLE V
Percentage Hit Rate of the Phonological Method for Identifying Chinese Surname

Validation Methods	Possible Chinese Surnames (N=266) Number of Surnames Verified
External Source	153
Panel	53
Direct Phone Interview	23
Total Validated Chinese Surnames (n)	229
Hit Rate (n/N) (229/266)	86.1%

DISCUSSION

The overall hit rate of 86.1% from the combined method indicated that the phonologically based surname search method is a valid and accurate technique for identifying individuals with a Chinese surname from the general population. Using the flowchart, both Chinese- and non-Chinese-speaking individuals can uti-

lize this method with ease, and the identification process is quick and low cost. Simple computer programs can be written to automate this process. Using phonological rules to identify other ethnic groups or languages is also feasible if the phonological structure of the ethnic group is distinct enough from the general population (e.g., Japanese). However, the strength of this study is not necessarily in repeating the

phonological method with other ethnic groups, but in using this method (unique for the Chinese language because of its distinct phonological structure) to identify Chinese surnames for research, survey or marketing purposes in many Western societies (e.g., France, Italy, the Netherlands, Germany, Switzerland).

This method does have some limitations. The 27 misidentifications were largely due to some surnames being common in other cultures (e.g., Lee, Kim, Bui, Guy, Duong, Ticio, Teoh). A few Western surnames also have the same phonological structure as Chinese surnames (e.g., King, Bean, Cook). Simple exception rules can be applied to minimize the extent of false positives. Names that are clearly Western names (e.g., Bean, Cook, Guy) should not be considered for selection using this method. Names that are common to other Asian cultures (e.g., Korean, Laotians, Vietnamese) should be selected with caution. These surnames are relatively easy to recognize and few.

False negatives might occur if the Chinese individual changed his/her surname as a result of cross-cultural marriages; or the Chinese individual was from a country (e.g., Vietnam, Indonesia) where a non-standard form of surname transcription method was used. Depending on the types of population being studied (e.g., elderly), the number of individuals who changed their surnames as a result of cross-cultural marriages might be insignificant. Finally, individuals who do not have a telephone or

have unlisted telephone numbers would not be identified using this method.

Despite these caveats, this method represents an important step forward in improving the cost effectiveness of sampling for studies specifically targeting the Chinese community. It avoids the serious pitfalls of reduced generalizability associated with commonly used methods like snowball sampling. Moreover, it can identify representative Chinese samples at a fraction of the cost of techniques such as random digit dialling.

For future research, a wider scope of search regions (e.g., Toronto, Vancouver) may be used to achieve a larger pool of Chinese surnames. Other major Chinese dialects should also be included to generate a more complete list of phonological rules and patterns. The phonologically based method developed to date is, nonetheless, a practical, user-friendly, valid, effective and efficient method of identifying the Chinese population from the general English-speaking population.

REFERENCES

1. Danesi M. The semiotic representation of "health" and "disease". In: Masi RM, Mensah L, McLeod KA (Eds.), *Health and Cultures: Exploring the Relationships, Policies, Professional Practice and Education*. Volume I. Oakville: Mosaic Press, 1993;243-54.
2. Zola IK. Oh where, oh where has ethnicity gone? In: Gelfand DE, Kutzik AJ (Eds.), *Ethnicity and Aging: Theory, Research, and Policy*. New York: Springer, 1979;66-80.
3. Chung RCY, Lin KM. Help-seeking behavior among Southeast Asian refugees. Special Issue: Asian-American mental health. *J Commun Psychol* 1994;22(2):109-20.
4. Hartge P, Cahill JJ, West D. Design and methods in a multi-center case-control interview study. *Am J Public Health* 1984;74(Jan):52-56.
5. Morton DJ, Stanford EP, Happersett CJ, Molgaard CA. Acculturation and functional impairment among older Chinese and Vietnamese in San Diego county, California. *J Cross Cultural Gerontol* 1992;7:151-76.
6. Raskin A, Chien CP, Lin KM. Elderly Chinese- and Caucasian-Americans compared on measures of psychic distress, somatic complaints and social competence. *Int J Geriatric Psychiatry* 1992;7(3):191-98.
7. Winkleby MA, Flora JA, Kraemer HC. A community-based heart disease intervention: Predictors of change. *Am J Public Health* 1994;84(5):767-72.
8. Masi R, Mensah L, McLeod KA. Introduction. In: Masi RM, Mensah L, McLeod KA (Eds.), *Health and Cultures: Exploring the Relationships, Policies, Professional Practice and Education*. Volume I. Oakville: Mosaic Press, 1993;3-10.
9. Samuel JT. *Visible Minorities in Canada: A Projection*. Toronto: Canadian Advertising Foundation, 1992.
10. Latimer E, Lundy M. The care of the dying: Multicultural influences. In: Masi RM, Mensah L, McLeod KA (Eds.), *Health and Cultures: Exploring the Relationships, Programs Services and Care*. Volume II. Oakville, Ontario: Mosaic Press, 1993;41-57.
11. Gordon MM. *Assimilation in American Life*. New York: Oxford University Press, 1964.
12. Wong SL. *A Chinese Syllabary Pronounced According to the Dialect of Canton*, 2nd ed. Hong Kong: "Zhong Hua" Publisher, 1995.
13. Avery P, Ehrlich S. *Teaching English as a Second Language: The Teaching of Pronunciation and Introduction for Teachers of English as a Second Language*. Toronto: Ministry of Citizenship and Culture, 1987.
14. Anonymous. *Standard Phonetic of Pai Chia Sing Chien Tzih Wen in English (hundred family names)*. Hong Kong: World Publishing Company, 1973.
15. Choi BCK, Hanley AJG, Holowaty EJ, Dale D. Use of surnames to identify individuals of Chinese ancestry. *Am J Epidemiol* 1993;138(9):723-34.

Received: August 18, 1999

Accepted: September 7, 2000