



# A Hermeneutical Approach of Gender Linguistic Materiality: Semiotic and Structural Categorisation of Gender in Hong Kong Cantonese

Julie Abbou, Angela Tse

## ► To cite this version:

Julie Abbou, Angela Tse. A Hermeneutical Approach of Gender Linguistic Materiality: Semiotic and Structural Categorisation of Gender in Hong Kong Cantonese. Julie Abbou; Fabienne H. Baider. Gender, Language and the Periphery: Grammatical and social gender from the margins, John Benjamins, pp.89-128, 2016, 9789027256690 ; 9789027266835. 10.1075/pbns.264.05abb . halshs-01955387

**HAL Id: halshs-01955387**

**<https://shs.hal.science/halshs-01955387>**

Submitted on 14 Dec 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A Hermeneutical Approach of Gender Linguistic Materiality: Semiotic and Structural Categorisation of Gender in Hong Kong Cantonese

Julie ABBOU\* & Angela TSE\*\*

\* University of Lorraine

\*\* Hong Kong Polytechnic University

## Abstract

While most of the Cantonese grammar textbooks draw on the assumption that gender is absent in this language, the fact, however, remains that the masculine/feminine relationship – and its consequential social assignation– is indeed expressed in Cantonese. This study explores how and where gender is indexed in the linguistic materiality of Cantonese. To serve such purpose, an interdisciplinary framework is built at the crossroads of linguistics and gender studies towards a hermeneutical re-definition of gender as a semiotic *form* structuring power relationships. This framework leads to an investigation of gender from a contrastive perspective. A corpus of translated sentences from English to Cantonese displays significant “gender shifts” in the inter-lingual space, which makes gender marking in Cantonese visible in different linguistic levels.

Keywords: Cantonese, English, Gender shift, Structure, Semiotics, Forms, Grammaticalisation, Contrastive, Interlingual dynamics

## 1. Introduction

This study aims at exploring possible answers to the following question: how does social gender linguistically work in Cantonese?<sup>1</sup> Most of the Cantonese grammar textbooks draw on the assumption that gender is absent in this language. If this were an invincible fact, the current survey could have been given up here. Yet, the linguistic fact remains that the masculine/feminine relationship is expressed in Cantonese and this chapter

explores how and where gender is indexed in the linguistic materiality of Cantonese. To address both these questions, a two-faceted approach is chosen. First, we opted for an interdisciplinary framework, which allows us a hermeneutical re-definition of gender at the crossroads of linguistics and gender studies, gender will then be defined as a semiotic *form* structuring power relationships. Second, we provide a distinct data collection method in order to make the gender marking in Cantonese in different linguistic levels visible.

In the following sections, the theoretical framework which has been built for this study is delineated with the intention of situating gender marking in Cantonese as a social, semantic and linguistic categorisation. Working on and with a language so-called genderless requires an understanding of gender as multi-dimensional dynamics which can be deployed simultaneously in language structures (i.e. gender as a formal opposition), in semiotic constructions (i.e. gender as taking part in the making of signification) and in social contexts (i.e. the material consequences of what makes of us men and women). Such an attempt to capture the different dimensions of gender poses questions about the definition of gender itself: when defining gender, are we discussing different but simultaneous dimensions of gender as an object? Or are we instead referring to different, concurrent, acceptations of the term *gender*? Postulating these diverse uses of gender as a distinct but interrelated categorisation process is noteworthy. Therefore, it is proposed in this study that an examination of *gender forms* can serve as a tool to highlight the categorisational power of gender. This proposed framework relies on gender studies, linguistic typology, hermeneutical semantics, and contrastive linguistics.

In order to study gender marking as a social, semantic and linguistic categorisation the methodological blueprint of this study is the following: to track in grammaticalisation the linguistic materialisation of the gender form as a semiotic category. The notion of periphery is then helpful here. If we consider a language structure as a set of potential and plastic categories (cf. Blache *et al.* 2014), then to involve the periphery of a category is to explore its plasticity. What matters in this research is how masculinity and femininity are *possibly* said, not how they are said i.e. to investigate the “likelihood” of gender in a genderless language. To do so we chose a contrastive and translational methodological approach in this study. The study of data obtained by inviting a sample of thirty bilingual speakers to translate sentences from English to Cantonese, two languages with different gender marking, made apparent instances where gender appears and disappears in an inter-lingual shift. By interlingual shift, we understand a contrastive move from a linguistic structure to another one. Therefore periphery is employed here as a methodological tool to understand linguistic categorisation. In order to make the most of the plasticity of a linguistic structure, written Cantonese has been selected to engage in this gender-

linguistic adventure because it is a mildly normalised and emerging variety, whose uses fluctates rather drastically.

This chapter thus pursues a dual objective: 1) it aims at formulating a theoretical proposition for gender analysis in genderless languages and 2) it provides a description of gender in an understudied language. To do so, the second section is a theoretical discussion of various gender analyses in linguistics with the aim of proposing our own reading of linguistic gender as a transversal element of linguistic structures, that is a hermeneutical reading of grammatical gender. Based on this hermeneutical reading, a methodology is then proposed, which involves an inter-level and contrastive approach, as well as the data used are presented (section 3). The fourth and the fifth sections consist in a presentation and a discussion of the linguistic materiality of gender as described in the literature about English and Cantonese, including sociolinguistic, lexical, morphosyntactical and metalinguistic elements. In section 6, the outcomes of the translation survey are presented and discussed. Although the gender marking in both two languages are of different syntactical embedment (i.e. gender being more grammaticalised in English), the results show that gender marking in Cantonese affects different linguistic dimensions and can be represented by the speakers' chosen strategies. The description of gender dissemination offers challenging yet insightful perspectives regarding our analytical tools, especially regarding the inflectional understanding of formal gender as well as the signifier/signified relation implied by a "morphology-first" approach of language. These perspectives constitute our conclusive remarks.

## 2. Interplay between grammatical, semantic and social features of the gender categorisation

We can identify two traditions in the study of gender and language. The first one, illustrated by Corbett's works (1991, 2013), considers gender as a syntactic *classification* of lexicon which relies on linguistics as its core field. This approach draws on the etymological meaning of gender, as <genus> (lat.) 'a kind, a sort'. Within this interpretation, the semantic dimension of gender is considered as secondary information, while its syntactical function is primary. The role of (linguistic) gender is therefore a contrastive opposition for parts-of-speech classification purposes, for example in French, grammatical gender allows a semantic distinction between *le livre* ('the book' [masc.]) and *la livre* ('the pound', [fem.]).

The second approach was launched a few decades ago, with author such as Cameron (1990, 1992), under the label of "gender and language studies", within which gender should be understood primarily as a social-semiotic construction either *reflected in* language or *constructed through* language. In this view, (social) gender is understood as a categorisation process which

allocates masculine and feminine features to individuals, a process in which language has definitely a role to play. Here, the semantic dimension of gender is primary and gender studies forms the foundation of the approach. To this regard, Chevalier notes:

For the ones who want to work on *gender* in its relationship with *gender*, there are two paths: either considering what gender studies owe to grammar, or studying what grammar owes to gender studies. (Chevalier 2013, 3)<sup>2</sup>

The aim of the current study, however, is not to find out which is indebted to which; rather, it is to explore what conceptual tools gender studies and linguistics may lend to each other, and what common ground they share to situate gender in language. Indeed, the two abovementioned standpoints (i.e. Corbett vs. Cameron) on gender draw on a distinction between what we may call an “intra-linguistic” and an “extra-linguistic” understanding of gender, that is, an interpretation of grammatical gender limited to the linguistic code vis-à-vis an understanding of grammatical gender involving social and linguistic relations. Pan and Tham (2007) evoke this distinction in terms of scales, labelling the intra-linguistic level as *micro-linguistics* and the extra-linguistic level *macro-linguistics*, the latter being understood as a contextualised view on language. In this chapter, therefore, it is argued that working about gender in language requires one to work from and towards micro- and macro-linguistics. Indeed, whatever approach is chosen, categorisation lies at the core of the matter.

Are linguistic gender and social gender shaped by an identical process of categorisation? To answer this question, we carried out an evaluation of the “intra-linguistic” dimension of gender at the outset. This intra-linguistic examination is then be complemented by a hermeneutical analysis to establish a language understanding which concurrently allows the explanation of the social, semantic and linguistic dimensions of gender and questions the linguistic structure with regards to the social structure. For this purpose, the insertion of a third dimension to the above intra- and extra-linguistic approaches, i.e. an *inter-linguistic level* approach, leads to a contrastive approach allowing the analysis of the linguistic materiality of gender as a social construction.

### *2.1. Gender and classification systems*

From a micro-linguistic point of view, gender is a matter of lexical classification. Grinevald (1999) argues that there is a continuum in the nominal classification, with at one end a lexical classification and at the other a grammatical classification; along the continuum gender falls into the grammatical pool, representing a sub-type of grammatical classification (see table 1):

**Table 1: Grinevald's Nominal Classifications**

| Lexical Classification | Classifiers | Grammatical Classification |
|------------------------|-------------|----------------------------|
| Measure Terms          |             | Noun Class                 |
| Class terms            |             | Gender                     |

However, Corbett (1991) and Kibort and Corbett (2008) consider gender as one of the two hypercategories of grammaticalised lexical classification besides classifiers, as shown in Table 2:

**Table 2: Corbett's Grammaticalised lexical Classifications**

| Gender                                   | Classifier                                 |
|--|--|
| Morphosyntactic feature (with agreement) | Morphosemantic feature (without agreement) |

This proposal is nowadays the mainstream in typology, indicating that noun classes are a sub-group of gender, with the specificity that their semantic references are non-sex-based. Pursuing his investigation on gender, Corbett, in a later chapter on gender in *The World Atlas of Language Structures Online* (2013), offers a classification of 257 languages according to the presence and absence of gender in them. He distinguishes between languages without gender, with one, two, three, four, and five or more genders, with instances like Nigerian Fula which embodies around twenty genders in some varieties. Referring to an example of a four-gender language called Lak, which is a Northeast Caucasian language, Corbett stratifies the gender classification in that language by the following classes: (i) male rationals, (ii) female rationals, (iii) other animates which interestingly contains some inanimate members, and (iv) a residue gender which also includes a few animates. In whatever way one chooses to label the hypercategories – *gender* for Corbett or *grammatical classification* for Grinevald, a range of formal features can be displayed. To this regard, Grinevald (1999) juxtaposes grammatical classification with lexical classification and eventually comes up with the summary below (see table 3):

**Table 3: Lexical Classification and Grammatical Classification according to Grinevald**

| Lexical Classification                                 | Classifier (a “hybrid” lexical-grammatical system) | Grammatical Classification (noun class + gender)                   |
|--|--|--|
| not all the nouns                                      |  | classify all nouns   |
| various amount of classes                              |  | little number of classes   |
| open system  |  | closed system  |
| not combined with other class                          |  | combination with other grammatical category (number, casuality...) |
| not assigned to a noun                                 |  | marked on the noun   |
| no agreement (once marked)                             |  | agreement system   |
| one noun = one class                                   |  | one noun can only be assigned to one class                         |
| differences from a speaker and register to another one |  | no speakers variation, no register variation                       |

Targeting a smaller-scale categorisation, Voisin (2014) alternatively highlights some differences between gender and noun classes based on three criteria:

- a. the semantic motivation (for gender, semantic motivation equals social gender, while semantic categorisation of noun class is much unclearer;
- b. the mark on the noun itself (i.e. gender is not marked on the noun, while noun classes are);
- c. the intrinsic feature of gender (i.e. a noun has one inherent gender, while in the noun class system a radical may carry different classes).

Despite the interesting and formal definition of gender provided by the above typological distinctions and criteria, nonetheless, some problems remain. First, in many languages (e.g. Romance languages) – if not all languages with a gendered classification built within their lexicons, gender works across linguistic levels and is thus often lexical as well as grammatical. In French, for instance, the noun *chauffeuse* (meaning ‘a low chair without arms’) carries a feminine grammatical mark, *-euse*, without being semantically motivated.<sup>3</sup> While being marked on the noun, the feminine gender embedded in *chauffeuse* clearly plays the role of a grammatical classifier. This hybrid dimension of gender can be commonly found in the French language because a number of very productive morphemes in French (e.g. *-ier*, *-eur*, *-teur*) are grammatically sensitive to gender while being semantically linked to inanimate referents and, most importantly, while not taking part in a gender contrast. Many words carry such morphemes without having a masculine *and* a feminine forms, like *saladier* (‘salad bowl’, [masc.]) / \**saladière* (\*[fem.]), *garçonnière* (‘shag

pad' [fem].) / \**garçonner* (\*[masc.]), *interrupteur* ('switch' [masc.]) / \**interruptrice* (\*[fem]). These examples show that the opposition between lexical and grammatical classification according to the mark of the noun does not provide a relevant criterion for a gender definition.

Consequently, in the light of the hybrid role of gender which encompasses both lexical and grammatical marking, a crucial issue regarding the problematic designation of sex as the reference of gender arises. This issue concerns the arbitrary and the motivated semantic reference of gender. Despite asserting that "there is always a semantic 'core' to the [gender] system, that is, there is an overlap between the nouns which take a particular set of agreements and some semantic features", Corbett (2013) slots languages into sex-based and non-sex-based gender systems though developing a method of categorisation which is tantamount to the Saussurean distinction between the motivated and the arbitrary dimensions of language. Citing Guinean Fula as an example, Corbett notes: "[Gender in Guinean Fula] has a clear semantic core; however, sex is not a part of it: nouns denoting human males and human females are found in the same gender. Other genders overlap with semantic categories to a greater or lesser degree" (*id.*). He states that the accordance between gender and a semantic category (sex, or rather social gender) is a matter of degree: "At one end of the scale, in sex-based systems gender may match the semantic category almost completely" (*ibid.*). For instance, in Tamil, which is spoken in Sri Lanka, it is almost always true to say that nouns denoting male humans are masculine, and vice versa. In Indo-European languages like French and Russian, however, while it is not entirely inaccurate to claim that nouns denoting males are typically masculine, a large proportion of masculine nouns do not denote males. Corbett therefore suggests that for gendered languages, gender can be attached to nouns according to semantic or formal features. Gender in English, for example, is assigned by semantic features, (i.e. the major distinction is based on the model *nurse/bishop* rather than *actor/actress*).

Further remarks on the above statements can also be made. First, the sex-based criterion reveals a classical pitfall of linguists when examining gender. Social construction of gender is then systematically erased, and the extra-linguistic world is rudimentarily perceived as biological, or natural. The biological metaphor has been a temptation in the linguistic field for centuries, implying a naturalist vision of gender and sex, in a clear opposition to a constructivist understanding of gender, such as delineated in gender studies, with Scott's (1988), among others. Paradoxically, most linguists remain blind to the light shed on the semantic dimension of gender by feminist studies. Linguistics studies based on a semantic conception of the extra-linguistic referent as natural, rather than a socially constructed set of significations, could only miss the paradigm of gender as developed in social sciences and humanities. From this point of view, beyond the



language structure only lies the wild world of Nature... Yet, as reiterated by Cameron, paralleling the proposition of Jakobson in 1971, “the concepts of ‘masculine’ and ‘feminine’ are infinitely detachable from anything having to do with ‘real’ sexual difference” (Cameron 1992, 82) due to the metaphorical function of language, while the sexualisation of gender, far from being *natural*, “was consciously constructed over centuries” (Burr 2012, 29) by grammarians, as demonstrated by Burr in her brilliant article about the concordance of sex and linguistic gender.

Second, the definition of gender becomes ambiguous when Corbett (2013) states that “there is always a semantic core to the gender system”, but at the same time acknowledges the existence of languages with 20 genders.

Lastly, among the 257 languages analysed by the researcher, most of them (i.e. 145 languages) are classified as not having a gender system. Of course, we cannot contradict this statement from a syntactical point of view. But, as soon as we understand gender as a more complex object, we realise that there is no society within which gender is not at stake, whatever the linguistic encoding of this distinction is. In other words, gender might be more or less grammaticalised, but it is always embedded in language. That is what Huddleston and Pullum (2008) argue when they propose that gender (which should be understood here as masculine and feminine relationship or opposition) is a matter of degree. Their *Cambridge Grammar* offers an interesting shift of view: gender is not *more or less semantic-based*, as stated by Corbett; rather, it is *more or less grammaticalised*. We then understand here that what is at stake is the definition of gender: whether it is primarily syntactical or primarily semantic (and, in fact, as this paper suggests, semiotic).

Corbett’s definition of gender is certainly an “intra-linguistic” one, including gender in the broader linguistic feature of the syntactical classification of lexicon. Nonetheless, this typological approach helps us define precisely the core of grammaticalisation and reveals the semantic segments where languages are grammaticalised.

This definition, however, points towards another tension. Erbaugh, using classifiers in the broad sense of nominal classification, writes:

languages with or without classifiers index very similar categories. The difference (...) reflects how *overtly* a particular language requires grammatical marking. Some languages mark categories such as number or gender more *overtly*.<sup>4</sup> But careful comparative testing finds no languages-caused difference in ability to think about quantity or sex. (Erbaugh 2002, 54-5)

Although extra-linguistic gender is once again understood as sex, Erbaugh in the above quotation confirms that the variation of gender across languages is a matter of grammaticalisation degree. This informs us about the process of formalisation, and of structuralisation of gender, as scalar.

Yet, as noted above, a mere syntactical approach to gender in language is not sufficient if we intend to address simultaneously the grammatical, social and semantic aspects of gender.

Indeed whatever the degree of formalisation gender displays in a language, gender categorisation is as much linked to the social structure as to the linguistic structure. The relationship between social structure, linguistic structure and discursive practices embedded in the semantic dimension of gender are addressed in the following section.

## *2.2. Towards a hermeneutical reading of grammatical gender*

If we want to confront both linguistic and social structures, we need to deal with the structuralist-only vision of gender. Indeed, although the institutionalisation of gender in language and society is beyond any doubt, the limitation of a structuralist paradigm prevents us from a complex comprehension of gender.

Structuralism has brought an indubitable and groundbreaking emphasis on relations to all academic fields which embrace it. This focus on relations rather than on terminology allows the abandonment of the realm of essence and substance. In linguistics, structuralism allows a contrastive comprehension of gender categories. However, this internal linguistic point of view implies a monosemic understanding of gender contrast. Once defined as a structuralised relation, the meaning of gender cannot be negotiated and the making of gender as a categorisation and as a power relationship is evicted. This is an internalist reading of language, which could be defined as functionalist:<sup>5</sup> language is a stable system made for communication, a homogeneous code that has to be reduced by linguists to its core in order to discover its internal laws. If such a mechanist approach in linguistics brings apparently a very inoffensive political stake (as we are speaking of sound's determination, instead of people's determination), the pitfall of structuralism lies in the logic-based comprehension of a language, making it a code which rids itself of its speakers. As a result, the linguistic *structure* passes from a modelisation proposal to a proclaimed reality which has to be fulfilled by the speakers.<sup>6</sup>

Given the fact that the social and the linguistic structures of gender are not equivalent, how should researches go beyond a mechanist and monosemic conception of gender opposition while maintaining the emphasis on relations with both society and language? In other words, how can gender be grasped in its simultaneous structural dimension (i.e. relational) and its semiotic dimension.

To answer this question, we believe that both structuralist and post-structuralist approaches to gender in social sciences and humanities are helpful in order to find a common ground shared by linguistic and extra-linguistic structuralisms: categorisation and meaning making.

In classic Chinese, the polyseme *bian* means ‘to discuss’, ‘to argue’ (as in 辯), as well as ‘to distinguish’, ‘to cut up’, or sometimes ‘to divide’ (as in 辨), using in both cases the ‘blade’ radical 刂 (Cheng 1997, 97). We find in this polyseme a hermeneutical understanding of language: meaning making is the categorisation itself. That is, the language structure is not a set of rules which lay out an extra-linguistic and pre-semantic world. Rather, language and discourse are “the place of social life and human affairs” (Rastier 2001, 111), the place where we negotiate realms of masculinity and femininity and we cut gender categories up. This is also a site where this attribution can be solidified through structuralisation, both in language and in society.

This is why we pose that a semiotic definition of gender allows the understanding of gender both in its social and linguistic dimensions, simultaneously. Such designation of gender as a primary semantic category meets the propositions by Violi’s (1987) and Scott’s (1988), from a linguistic and a historical perspective respectively. To verbalise gender, therefore, is then a way to signify gender in reality. In Butler’s words, the power of categorising gender takes part in the making of intelligibility (1990). Here, Butler echoes Whorf:

We cut nature up, organize it into concepts, and ascribe significances as we do. (...) we cannot talk at all except by subscribing to the organization and classification of data which the agreement decrees. (Whorf 1940, 212)

This semiotic approach challenges the canonical Saussurean linguistics with regards to the *langue/parole* (or discourse) opposition and to the arbitrary/motivated opposition.

Within this positioning, we give significance to the world through gender and we structuralise this categorisation in order to institutionalise and grammaticalise it. The resonance between Whorf and Butler puts forward the following statement: the performative categorisation of gender not only happens in discourse but also in the language structure itself. In this view, structure (*langue*) and discourse are not in a causal relationship (whatever being the cause or the consequence) but rather in a co-construction dynamic in the light of a power-relationship and of a normative process. Structure is then a solidification of one of the possible categorisation proposals.

The hermeneutical (interpretative) framework also invalidates the arbitrary/motivated opposition based on the refusal of a pre-semantic extra-linguistic world. From the hermeneutic point of view, linguistic gender is never motivated by a pre-linguistic gender or a biological sex, and the arbitrary dimension of gender is the result of gender formalisation and institutionalisation.

Rather, gender semantics is embodied in a language structure. Diverting from the notion that language is a mere representation of reality, one can

investigate how gender takes linguistic materiality in language and the manner which gender is linguistically institutionalised, normalised, structuralised, and included in so-called genderless languages. Again, we need to reverse our standpoint from a syntactical point of departure into a semantic one; informed by micro-linguistics, we need to turn towards macro-linguistics.

By referring to Rastier (2001), we can therefore consider language as a repository of semantic forms and contents. Paralleling the fact that gender varies through space and time, linguistic arrangement of gender varies from one language to another. Linguistic marking of “doing gender” (West and Zimmerman 1987) occur in every language, and this is precisely how Rastier defines what a linguistic form is: it is transposable in nature (Rastier *ibid.*), and this implies that if we want to catch a glimpse of the *gender form* in a language, we need to look at how gender forms are transposed from a language to another. What we are suggesting is therefore not to look at what is contained in the feminine or the masculine category, but to look at how sensitive the distinctive process is in different contexts and languages.

However, gender in its social-semiotic understanding bears another form of specificity. Not only is gender a categorisation, it is also a *bi*-categorisation. He-Yin Zhen, a Chinese feminist in the early twentieth century, recognises in this *bi*-categorisation the fundamental process of gender (Liu, Karl and Ko 2013). She uses the concept of *nannü youbie* to shed light on the distinction as a founding act of categorising and gendering the world. She considers “the marking of the difference, the act of making something distinct” (Liu *et al.* 2013, 53) as prior to the distinction itself. According to Zhen, who wrote it in 1901, gender is primarily a categorising process moulding power-relationships. This categorising power is deployed similarly in the social assignation to masculinity and femininity and in the linguistic structure of gender. Butler speaks of a *matrix of intelligibility*: “an open system of signs by which intelligibility is insistently created and contested” (1990, 184).

To summarize, in the two analytical sections above, a linguistic portrayal of gender has been drawn grounded in five dimensions: (1) scalarity of gender grammaticalisation, (2) semiotic dimension as the primary dimension of gender, (3) structuralisation as a solidification of the gender categorisation process, (4) definition of gender as a form, that is, a transposable relationship, and (5) Zhen’s *nannü youbie*, the dichotomous distinction as a power-process founding the gender categories.

In order to take into account these different dimensions of gender in our analysis, a dedicated methodology has to be designed.

### **3. Methodological consequences: An inter-level and inter-lingual approach**

The conclusions drawn from the section leads to the abandonment of the notion of *language without gender* proposed by Hellinger and Bussmann (2001, 2002, 2003), when gender is grammaticalised in some languages while in others it is only a semantic feature. Such a view relies on an understanding of a dichotomy between arbitrary and motivated gender. But what would be a semantically motivated content without a linguistic form? Instead, we should explore the interaction between different linguistic levels, such as morphology, lexicon, syntax and discourse, where gender functions from a semiotic perspective. We suggest to combine in the analysis all the linguistic levels altogether. We understand that Hellinger and Bussmann distinguish between languages with and without grammatical categories of gender, but we assume in this current study that this dichotomous distinction, i.e. syntax vs. other linguistic levels, is not useful for the development of a critical methodology of gender and of language studies beyond the level-by-level approach. On the contrary, the interplay between ideology and linguistic materiality makes language a space of reality negotiation (see Angenot 1982 for a detailed theorisation of this interplay). This positioning requires a consideration of the institutionalisation of gender in language as an inter-linguistic levels continuum: linguistic gender is primarily a dichotomous semiotic notion (Violi 1987; Karl *et al.* 2013) *more or less* grammaticalised/formalised (Huddleston and Pullum 2008) depending on each language transposition (Rastier 2001).

Furthermore, these notions of transposability and scalarity on a continuum lead us to a contrastive approach. The articulation between intra-linguistic and extra-linguistic understandings of gender, as two simultaneous dimensions of the categorial processes of masculinity and femininity, has been discussed above; such background has paved the way for the addition of the inter-lingual dimension to the two-faceted analytical framework. By transposing gender from one language to another, the potential variations of the masculine-feminine relationship can be revealed. This is an attempt to work on the plasticity of gender forms across languages.

By comparing two “non-gendered” (or covert-gendered) but distant languages, i.e. English and Cantonese, we formulate the hypothesis that it is possible to describe *in relation* the linguistic embedment of the semantics of gender in Cantonese. English in the current study is chosen as a contrastive language because of its well-described literature on gender (see section 4 below), but also due to the postcolonial context in Hong Kong, in which both Cantonese and English are commonly used. Indeed, the interlingual shift between the two languages is a common practice for Cantonese speakers.

In various literature, the comparison of gender role in different languages mainly concern the overt and covert dimensions of a category. The overt categories concern grammaticalised features, while the covert categories

(the cryptotypes) are not explicitly coded in grammar or lexicon, but rather emergent only in syntax and various usage patterns in the language (see Moser 1997 for an example of the use of these categories to describe gender in Mandarin). According to our theoretical framework, we use the notion of more or less grammaticalised categorisation (therefore, more or less marked signs) to outline the structuralisation at stake rather than the implicit or explicit character of gender.

We conducted this contrastive approach conducted first on the linguistic description of gender in grammar and linguistic studies about English and Cantonese. However, we were careful during the comparison, because of the paradox it entails: to build common tools in order to measure differences. This paradox contains a risk, already outlined in Whorf's work:

The very natural tendency to use terms derived from traditional grammar, like verb, noun, adjective, passive voice, in describing languages outside of Indo-European is fraught with grave possibilities of misunderstanding. (Whorf 1937, 87)

This risk is actually a risk of grammatical colonialism, by using the tools of Western linguistics to describe non-Western languages. On the other hand, however, a globalised academic world with hybridised linguistic analyses renders it difficult and almost impossible to segregate “pure” Western language evaluation from the non-Western ones without re-essentialising the scholarly traditions as endogenous and homogeneous. Rather, the contrastive approach should lead to an acknowledgement of the domination interplaying between languages and between linguistic knowledge. From this perspective, we can question our analytical framework instead of adopting it at any cost. The inter-level approach serves as a safeguard in the contrastive work. Indeed, looking at different linguistic levels simultaneously reduces the risk to focus on already described phenomena only.

To complete this comparison, *gender shifts* from a language to another one, have been observed through a translational process. Given the categorisation of gender being different in English and Cantonese, the inter-lingual dynamics (defined as the move from a language to another one) outline the gender form variation in each of the two languages. Thirty bilingual speakers were invited to translate twenty English sentences into written Cantonese. When translating, participants needed to transpose gender information from one language to another. For this purpose, they sometimes disambiguated gender information, that is, they might either have added gender information or simply shunned it by erasing gender details. This translational moves help capture the different linguistic designs of gender as a semantic category, in English and Cantonese. This resonates with Lado's Transfer Theory (1957) which seeks to describe how speakers transfer habits from a language structure to another language structure.

In order to address linguistic plasticity, we also need to scrutinise the descriptions of standard languages (or grammars) as non-standard productions. For this purpose, we need to find linguistic productions which are less subject to normative pressure. Written Cantonese is a relatively (re-)emerging variety without standardisation, without legal or cultural recognition; this language embodies therefore such a characteristic of under-normalisation. The advantages of working on Cantonese are presented further in the dedicated section below.

Lastly, there are two points which require clarification. First, the objective of this study is not to evaluate or score languages according to sexism. Whether gender is a cryptotype or a grammaticalised linguistic item does not make it easier to deconstruct or to subvert. Rather, it is about identifying the variety of the forms gender can endorse across languages. Second, we are not proving the universality of gender. What is of interest, we believe, is not the outline of the universally deep-rooted patterns by means of variation. On the contrary, the variation reveals the construction of gender as a power mechanism penetrating every dimension of social life. Following the line of Whorf, we seek to “experience an interruption of phenomena intherto held universal” in order to see the emergence a whole new order of significance (Whorf, quoted in Pan and Tam 2007).

Consequently, this study relies on two sets of data. First, the literature about gender in Cantonese and English languages is discussed (sections 4 and 5), especially grammar textbooks and studies in syntax. This critical review of gender analyses within these two linguistic structures reveals the criteria and theoretical realms used for specific gender analyses in English and Cantonese. It provides as well an insight of the debates framing these analyses. Second, a corpus-driven analysis in section 6 is presented based on a survey implying an English-Cantonese translation task. This part enables the confrontation of the syntactical literature with actual uses.

#### **4. Linguistic gender in English**

While the literature on Gender and Language is predominantly written in English, the scholarship concerning linguistic gender in English language *structure* remains scattered (see Silverstein 1985; Baron 1986; Corbett 1991), not to mention a feminist approach of the topic (Cameron 1992; Hellinger and Bussmann 2001; Romaine 2001; Motschenbacher 2010; Pauwels 2010) compared with the multifarious works in sociolinguistics and discourse analysis. We then believe there is a need to merge more closely the study of the linguistic spaces and the study of the categorical power of gender in English.

Corbett (1991) puts forward three spaces in terms of “levels”:

- i. semantic level: terms with a gender connotation or overtone: *bishop, nurse, whore, handsome, cute*, or *wallet/purse*. Here, gender is secondary semantic information;
- ii. lexical level: human-referred terms with a gender denotation: *boy, mother, uncle, wife* and composition. Unsurprisingly, most of these terms imply a relationship, while gender is the relation between masculine and feminine. Here, gender is primary semantic information;
- iii. morphosyntactical level: agreement phenomena such as pronominal anaphoric control: *he, she* and some derivations: *actor/actress*).

Silverstein (1985) moves away from a linguistic level analysis and proposes an analysis by “dimensions” based on the different dimensions of language:

- i. structural dimension;
- ii. pragmatic dimension;
- iii. ideological dimension.

Motschenbacher (2010, 64f), mixing a level and dimension approaches, suggests the following classification:

- i. lexical gender;
- ii. social gender;
- iii. grammatical gender;
- iv. referential gender.

These different classifications of gender of the English language which incorporate social, or ideological, and structural dimensions are in favour of a hermeneutical reading of language. We argue here that an analysis of gender and language and gender in language will be benefitted if such a reading from a linguistic and feminist/gender perspective is adopted.

One can read in the *Cambridge Grammar of the English Language* that “the basis of the distinction between different sorts of name is not necessarily the sex of the referent” (Huddleston and Pullum 2008), as illustrated by the comparison of *he/she/it* with *who/which*, within gender functions as a grammatical classifier in English. Unlike “gendered languages”, such as Romance languages, the grammatical role of gender in English is not to classify the entire lexicon but to distinguish between the animate and the inanimate, and the above demonstrates that gender in English encompasses a grammatical function. It also implies that, in the English language, the human/non-human feature is a primary signified, while gender is a secondary signified. The discussion above is presented in table 4, within which gender is embedded in broader grammaticalised semantic processes:

**Table 4: Morphosyntactical classifiers in English with semantic roles**

| Singular        |                       | Plural                 |
|-----------------|-----------------------|------------------------|
| Human (animate) | Non-human (inanimate) | <i>they/them/their</i> |



|                                |                                |                               |  |
|--------------------------------|--------------------------------|-------------------------------|--|
| Masculine<br><i>he/him/his</i> | Feminine<br><i>she/her/her</i> | (Neutral)<br><i>it/it/its</i> |  |
|--------------------------------|--------------------------------|-------------------------------|--|

Well-known diachronic elements reveal that this grammaticality of gender in English are residual tracks of a past stronger grammaticality:

Old English had three grammatical genders – masculine, feminine and neuter – and all inanimate nouns belonged to one of the three classes, sometimes for morphological reasons but often for no obvious reason. (Curzan 2003, 12)

The famous “werewolf example” of *wereman* and *wifman* (see Cameron 1985) is a lexical mark of this past grammaticalisation.

## 5. Linguistic gender in Cantonese

Since the late 1990s, the academic conversation on gender issues in Cantonese as a language has been expanding. Similar to the literature about English, studies on Cantonese are mostly dedicated to the representation of women and men in different contexts: business (Schnurr 2008, 2010), ad landscape (Fung 2006; Kang 2008), textbooks (Law and Chan 2004; Lee and Collins 2008; Yang 2011) or on the Internet (Kang and Chen 2012). To a lesser extent, there are also some works on the linguistic practices with regard to speakers’ gender (Chan 1999). It is not surprising that languages whose gender is little grammaticalised are rarely described from a linguistic point of view and are rather studied as an issue pertinent to sociolinguistics and discourse analysis. However, as one of the assumptions given in our introduction, the linguistic forms are also a place for gender embedment. Some grammar textbooks of Cantonese, therefore, inform how parts of speech are associated with gender. In this section, we relies heavily on a major grammar reference of Cantonese, *Cantonese: a comprehensive grammar* by Matthews and Yip (1994). Before comparing the embedment in English to that of Cantonese, some sociolinguistic elements of the Cantonese language should be introduced in order to contextualise the upcoming linguistic analysis.

### 5.1. Sociolinguistic elements of Cantonese

With more than 62 million speakers in the world (Lewis, Simons and Fennig 2014), Cantonese (also known as Yue or Gwongzau Waa) is mostly spoken in the south-eastern part of China, e.g. Hong Kong, Guangzhou, Guangxi and Macau, and the use of Cantonese in Hong Kong is chosen to be our exclusive focus.

Ceded by China to the United Kingdom in the late nineteenth century for almost one hundred years, Hong Kong's history in the twentieth century was quite separated from Mainland China. 1997, the year of Hong Kong's handover to China, marked the commencement of a fifty-year transitional period before Hong Kong fully belongs to the People's Republic of China (PRC). Hong Kong is therefore nowadays a transitional, postcolonial and multilingual city.

According to the Census and Statistics Department of Hong Kong government (2012), Cantonese has been the usually spoken language of 89% of the Hong Kong population for the past 20 years. This is one of the three official languages adopted by the government of Hong Kong besides English and Mandarin.

Cantonese in Hong Kong is “clearly the most omnipresent [language] used in Hong Kong schools, media, government, business, and daily life” (Snow 2004, 1). Yet, Hong Kong “has not had an organisation formally assigned the job of monitoring or directing the development of Cantonese speech” (Bauer 1988, 285), which is still true in the twenty-first century. The language planning in Hong Kong consists mainly of maintaining a high standard of English and/or promoting Mandarin literacy, but nowhere it is on the agenda to promote Cantonese – it only functions as the major medium of instruction in some, but not all, primary and secondary schools. Cantonese is thus a non-institutionalised major language used by citizens in the Asian cosmopolitan city.

Such a situation creates a singular relation between the languages *in presentia*. Poon (2004) speaks of a triglossic situation, within which English is the high-standard variety, while Mandarin and Cantonese are the low-standard varieties. Snow (2004) describes the Hong Kong triglossia as follows:

- English
  - for international communication purposes
  - value = prestigious (vs. Cantonese and Mandarin)
- Mandarin
  - for most of the Chinese written communication purposes
  - value = official, serious (vs. Cantonese)
- Cantonese
  - for most oral interactions in Hong Kong and informal written communication
  - value = informal

This informal status of Cantonese is the reason of the absence of standardisation. Yet, some linguistic descriptions are available, where we can find identification of gender.

### 5.2. *Lexicon: relational terms and keys*

Lexicon is the most noticeable site where gender is marked in Cantonese. There are of course several relational pairs primarily referring to gender:

- (1) a. 男 *naam4*<sup>7</sup> 'man'
- b. 女 *neoi5* 'woman'
- c. 仔 *zai2* 'son'
- d. 女 *neoi5* 'daughter'

Besides, “gender keys” such as 女 *neoi* and 男 *naam* can be used as a compositional prefix:

- (2) a. 女醫生 *neoi-jisang* 'female doctor'<sup>8</sup>
- b. 男醫生 *naam-jisang* 'male doctor'
- c. 女老師 *neoi-lousi* 'female teacher'
- d. 男老師 *naam-lousi* 'male teacher'

Since the gender keys encode primarily gender information (man/woman [affix]), they can also be used as lexical items by themselves (man/woman [word]).

On a lexical level, Cantonese is then very similar to English, with primary gendered terms referring to relationship and kinship, genderable terms by the addition of a gender prefix (e.g. ‘female doctor’ 女醫生), and a major part of the lexicon being unmarked.

At the crossroad of lexicon and ideology, it is of interest to note that the egalitarian rhetoric and ideology in post-1949 China fought against feminine marks both in language and behaviour. For example, the oppositional pair 妻子/丈夫 *qīzi/zhàngfu* ‘wife / husband’ was replaced by a new term: 愛人 *àirén* ‘spouse’ or ‘lover’ in Mainland China. However, such linguistic interventions have never happened in Hong Kong. Rather, there has been a subversive reaction against this neutralisation process. For example, the term 同志 *tóngzhi* ‘comrade’ which was used during the Cultural Revolution is today commonly used in Hong Kong and Taiwan for auto-reference in gay communities, and this new and subverted meaning has in turn impacted on its original use in Mainland China (Wong 2005; Kane 2006).

### 5.3. *Morphology: sentence particles*

The lexical classification in Cantonese is of high complexity: there are quantifiers, measure words and classifiers of different kinds (e.g. sortal classifiers, general classifiers, verbal classifiers, nominal classifiers).

According to Erbaugh (2002), these classifiers were originally semantic-based, but are nowadays formalised and endorsed several functions, such as classification, individuation, referentiality and relation (see Erbaugh 2002 for a detailed account of classifiers in Cantonese). As stated by Grinevald (1999) and Corbett (1991), it is rare to find languages presenting both a classifier system and a gender/noun class system. However, as shown above, the total distribution of the lexicon within a gender system is only the extreme of gender grammaticalisation. A presentation of some sentence particles in Cantonese reveals that some classifiers have to do with gender. Cantonese attaches sentence particles, which are nominal and adjectival suffixes, to gender-neutral or meaning-making terms to specify generic information or represent gender identities. Here are a few examples concerning gender excerpted from the aforementioned grammar volume by Matthews and Yip (1994):

- (3) a. 佬 *lou2*            a colloquial suffix for masculine referents.  
Ex.: 鬼佬 *gwai-lou* ‘western male foreigner’
- b. 婆 *po4*            a colloquial suffix for feminine referents.  
Ex.: 鬼婆 *gwai-po* ‘western female foreigner’
- c. 仔 *zai2*            a suffix indicating smallness for masculine or generic terms.  
Ex.: 香港仔 *hoenggong-zai* ‘Hong Kong people’ (generic meaning)  
花仔 *faa-zai* ‘flower boy in wedding’ (masculine meaning)
- d. 女 *nei5*            a suffix indicating smallness for feminine terms.  
Ex.: 花女 *fa-neoi* ‘flower girl in wedding’

Regarding gender, there is thus a morphologisation of lexicon in Cantonese. According to Grinevald, we are here facing a classifier system, within which gender can be added. Indeed, there are a number of sentence particles in this open system which are not assigned to nouns, and each noun may bear different sentence particles. Gender is therefore partially a grammatical object along the lexical axis of the continuum.

However, this statement has to be contextualised. The morphological and the lexical levels of Cantonese, which is predominantly a non-inflectional language, can easily overlap. To add the female key 女 before a character can either be a composition or an inflection. The written system guides us, according to the number of compounds:

- in 女仔 *nei-zai* ‘girl’ the feminine mark 女 is a compound, that is a lexical item, taking part in a compositional process

- in 嫁 *gaa3* ‘girl marrying someone’, the feminine part is included in the compound and therefore falls under an inflectional process, in the realm of morphology.

Moser addresses this issue in Mandarin when he discusses the limit of a compound:

the principle of affixation functions differently in Chinese than in English. (...) Does the prefix *nü* ‘female’ in compounds like *nüshen* result in these being perceived psycholinguistically as analytical two-word phrases corresponding to the English ‘female god’, or as more tightly-bound chunks like ‘goddess’? (Moser 1997, 16)

The number of compound remains the sole criterion to describe lexicalisation or grammaticalisation.

#### *5.4. Syntax: the written pronominal system and its evolution*

If the morphological realm gives more room to gender in Cantonese than in English, the syntactical space reverses the trend. Due to its isolating tendency, the Cantonese pronoun system does not adopt inflections, neither for gender, case nor animated feature (contrarily to English): “The pronouns have a single form for subject or object. There are no separate forms for the genitive (possessive) pronoun” (Matthews and Yip 1994, 80). Inflection, if it happens in Cantonese, concerns number only.

However, according to Matthews and Yip (*id.*), despite this assertion, we can read in the Cantonese grammar that the second-person singular Cantonese pronoun *nei5* (‘you’) could vary in gender when written. The most common character in use remains the masculine form 你 for referring to both genders. However, 你 and 妳 (the second-person singular pronouns of male and female respectively) are commonly used in written Cantonese. The masculine/generic radical 亻 is replaced by the feminine radical 女, changing the pronoun from 你 to 妳. The character 妳 thus specifies the feminine gender of the second-person singular pronoun.

Yip and Matthews (*ibid.*) also note a very rare phenomenon that a gender variation could also appear in the written form of the third-person singular pronoun.

##### *5.4.1. Linguistic version of the tradition/modernisation opposition*

In order to explain such a graphical variation in an isolating language within the written system is largely minor in regards to the spoken system, a socio-syntactical discussion about the evolution of Mandarin and its meta-discourses is needed. In the course of the twentieth century, while experiencing some (violent) colonial encounters with the West, China was concurrently infused with a range of western ideologies such as socialism,

liberalism, anarchism, evangelism, social Darwinism (see Cheng 1997; and Karl *et al.* 2013 for a more detailed account of this ideological sudden burst). These new upheavals happened on a controversial ground, particularly within the youth who were fighting against imperialism. The political divide was then weaved around the tension between modernity and tradition, both of which signify a progressive stance in a new era and the past conservative empire respectively.

This political embodiment was transposed and translated into the linguistic realm. On the modern side, Mandarin was strongly depreciated because of its isolating features. These features were considered less evolved and less civilised than an inflectional language. The progressive New Youth movement perceived it as an impediment to modernisation: “Chinese intellectuals agonised over the nation’s sad fate and quite a few became convinced that one major, if not the sole cause, which prevented China from modernising itself was its language” (Lee 2008, 69). Simultaneously, the high complexity of its written system, which was seen as a sophisticated symbol of the Chinese tradition, was treasured and maintained by the conservative wing. The writing simplification reform originated from this debate. Starting with the New Youth movement in the early twentieth century, with the support from intellectuals such as Ba Jin (Snow 2004), Chinese language simplification was achieved by the Communist Party (Bourgeois 1978) in the 1960s, and the simplified form was added an emphasis on cultural and literacy access.

Here is an example comparing traditional and simplified characters:

- |                      |                          |
|----------------------|--------------------------|
| (4) Traditional:     | 我到過美國                    |
| Simplified:          | 我到过美国                    |
| Mandarin Pinyin:     | <i>wǒ dàoguó měi guó</i> |
| English translation: | ‘I have been to America’ |

It is important to note that Hong Kong and Taiwan do not share the communist history of Mainland China and have been using traditional characters in reading and writing.<sup>9</sup> Under the rule of the British government, the insistence on the use of Cantonese and traditional Chinese characters instead of Mandarin and simplified characters was even supported as a tool against China. The use of traditional characters may therefore not be an instrument to maintain Chinese tradition (as there is no such thing as a high variety of Cantonese in Hong Kong) but rather a colonial strategy and an identity marker.

#### 5.4.2. *The Westernisation debate*

Beyond the simplification of the writing system, the traditional complexity and the isolating feature of the Chinese language led in the 1930s some intellectuals to pursue a modernisation of their mother tongue through

westernisation: “A good [Chinese] vernacular language is a westernised one” (Fu Sinian, quoted in Peyraube 2001). The creation of a written feminine pronoun in this frame supported a will to westernise, or “civilised”, language and subsequently society by using the narrative of the liberation of women and identifying Barbary with the masculine domination as a measure.<sup>10</sup> This colonial narrative still exists in recent years to justify the western military intervention in Afghanistan.

Another hypothesis of the appearance of this new written gendered pronoun, which is also linked to a Western influence, is the presence of Jesuits missionaries in China who described the inflectional grammar of Chinese using the analytical tools traditionally for Western linguistics. Whatever such gender pronoun invention came from Chinese intellectuals, Jesuit missionaries or a Western linguistic and colonial ideology, this phenomenon is consensually known as “westernisation”.

However, Peyraube (1999) makes a relevant point on westernisation of the Chinese language. If the Chinese grammars since the *Ma shi wen tong* (1898) were indeed influenced by Western grammars (and among them the Port-Royal grammar), this influence concerned Chinese linguistics rather than the Chinese language itself. Through a study of the history of language, Peyraube proves that the so-called westernised forms were attested long before any regular contacts with the West: “the influence of Western languages on the emergence of new grammatical structures in Chinese from the second half of the nineteenth century was nearly inexistent” (2001, 25). The Westernisation of Chinese language was therefore more a language planning project than a syntactical evolution due to linguistic contacts. Yang Zijian has a similar observation:

Chinese intellectuals, mindful to their mission to empower China with tools that will take the country to modernization, signified by military might and economic strength, have wasted no time in taking their cue from the West. [...] they looked hard] to establish a Chinese linguistics system based on the Western model. (Yang 2007, 4)

Starting with the *Ma shi Wen tong*,

Indo-European linguistic theories [were] introduced in [Chinese Linguistics], in particular two significant grammatical categories that have no equivalent in Chinese: word and sentence, not to mention a string of other categories such as gender, number case, tense, aspect, voice, person, mood, noun, verb, adjective, pronoun, adverb. (Yang *id.*)

Westernisation, as a language planning rather than a linguistic evolution, therefore resulted in gender marking in Mandarin. Kane notes that in the first half of the twentieth century, there were “attempts to introduce words

such as *yi* (for ‘her’) or *tuō* (for ‘it’). But these attempts failed” (Kane 2006, 107). Such effort to visibilise women corresponded to a gender ideology which would be replaced in the second half of the twentieth century by a neutralisation strategy, following the Maoist egalitarian collectivist ideology of making “the woman a worker as others”, as the motto went.

Supplying further illustrations of this attempt, Peyraube dates the appearance of distinctive characters for the third person singular pronoun from 1917: “It sounds likely that the graphic distinction between masculine 他 *taa1* ‘he’, feminine 她 *taa1* ‘she’ and neutral 它 *taa1* ‘it’ is dated from 1917, as suggested by Wang Li (1947, 368)” (2001, 23). Moser dates it from the May Fourth Movement, in 1919. It remains unclear if the change was introduced for egalitarian purposes (understanding the generic as primary masculine and therefore a need for women visibilisation) or was a will to strengthen the gender order by remasculinising the masculine in contrast to the feminine (shifting thus from a gender-inclusive masculine to a gender-exclusive one). The pronominal gender variation in Chinese is therefore a relatively new morphosyntactical feature, specific to the written system.

As for Cantonese, Egerod (1982, 803) writes that the generic morpheme *khoey4* (or *keoi5*; a third-person singular pronoun) – but not the character – has a history spanning about three thousand years. Concerning the writing system, the third-person pronoun is precisely one of the differences Bauer (1988) identifies between written Cantonese and written Mandarin. Among a classification of the formation process of Cantonese characters, the third-person pronoun resulted from a character creation to represent a Cantonese morpheme, as shown in Table 5, quoted from Bauer (1988):

**Table 5: Third person pronoun in Written and Spoken Cantonese (Bauer 1988)**

| Cantonese Character | Cantonese Pronunciation | English Gloss | Standard Chinese Character Equivalent | Putonghua Pronunciation |
|---------------------|-------------------------|---------------|---------------------------------------|-------------------------|
| 佢                   | khoey.4                 | ‘h’           | 他                                     | ta.1                    |

So, if this Cantonese morpheme has a three thousand-year history, if written Cantonese is a relatively recent phenomenon, and if gendered pronouns appeared in written Chinese not until the early twentieth century, one can reasonably form the hypothesis that the gendering of the written pronominal system of Cantonese is posterior to the gendering of the Mandarin pronominal system. Therefore, what we are dealing with in this project is a



recent gender distinction on a language evolution scale which is not necessarily a western importation and is still little spread.

More recently, Moser observes the appearance of a gendered second-person pronoun in Hong Kong and Taiwan. She condemns it as a de-genericisation, or rather a masculinisation of the generic pronoun:

More and more people (mainly in Taiwan and Hong Kong) are beginning to use a relatively new character for the second-person singular 妳 (ni, 'you [female]') with the female radical (女) which is in opposition to the formerly gender-neutral character 你 with the 'person' radical (你). Thus, the exact same gender inequality has been expanded to this case, as well. The character 你 now refers to either males or those of unspecified gender (or [plural]), whereas the character 妳 can only refer to females. (Moser 1997, 12)

Hence, a historicised reading of syntax shows that English and Cantonese present reverse tendencies regarding gender grammaticalisation, with a late and light genderisation of grammar in Cantonese (especially pronominal system) and a gender loss in the English syntax. After a brief review of gender in English and Cantonese grammars, we can now turn towards gender marking in use in the light of inter-lingual shifts.

## **6. Gender translation: an English/Cantonese comparative survey**

### *6.1. Corpus presentation*

We have seen above that gender is present in different linguistic levels in Cantonese and English, with an opposite diachronic evolution. To pursue the investigation, thirty bilingual respondents who are native Cantonese speakers were invited to translate twenty sentences from written English to written Cantonese. As mentioned in 3, the translation survey was designed to observe the gender form in Cantonese. For this purpose, the English sentences being translated included various types of human-referred terms (nouns, pronouns, anaphora, sentence subjects or objects, genitives, singular and plural nouns, etc.), non-human animate referred terms, generic terms, etc. Deictics were also included to test their influence on participants' understanding of gender elements in Cantonese (e.g. *I'm an actress*). In order not to explicitly direct participants' attention to gender, decoy sentences without any gender information were included and the survey was plainly presented to the respondents as a study about English-Cantonese translation and sentence length. Prior to the presentation of the research

outcomes, it is necessary to rationalise our choice of researching into the written form of Cantonese.

## 6.2. *Written Cantonese: Contextualisation*<sup>11</sup>

The spoken/written relation in Hong Kong is of high complexity. Indeed, most Cantonese speakers speak Cantonese, but Standard Chinese (i.e. the written form of Mandarin) is the “written variety of Chinese taught in Hong Kong schools and most often used in Hong Kong society” (Snow 2004, 2). However, if a Cantonese speaker simply reads aloud a sentence of Standard Chinese, communication could be impeded: some changes in grammar and wording are needed since Cantonese and Mandarin are two languages with considerable differences in terms of syntax and lexicon without shared understanding. This is therefore a very distinctive situation, with speakers who mostly read and write in a language other than the one they actually speak. In other words, the primary written language of Cantonese-speaking Hong Kong people is not the primary spoken language of the community. Based on this situation, a “biliterate-trilingual policy” was launched in 1997. “Biliterate” stands for English and Standard Chinese literacy, while “trilingual” stands for English, Mandarin and Cantonese proficiency (Poon 2004).

Yet, written Cantonese exists. According to Bauer, “written Cantonese and written Standard Chinese are two related but mutually-unintelligible written Chinese languages – just as the spoken languages (or dialects) are clearly-related but mutually-unintelligible” (1988, 249). They “coexist in a complex relationship that is both symbiotic (e.g., [Cantonese] borrows standard characters when necessary) yet separate: it gets its streak of independent development from spoken Cantonese which is unintelligible with Putonghua [i.e. Mandarin]” (*op. cit.*, 277). To write in Cantonese is then an attempt to bridge the lexical gap between Cantonese speech and written Standard Chinese.

Written Cantonese is in fact a re-emergent variety. Snow observes, “since the late Ming dynasty (1368-1644), hundreds of inexpensive editions of books of verse using Cantonese have been published” (2004, 6). Today, one witnesses an important spread of written Cantonese since the 1970s, going hand in hand with an economic and cultural transition, characterised by the growth of the mass media and printing technology (Snow 2004) and the emergence of a Hongkongese identity (Ku 2004). In 1988, Bauer already mentioned a strong development of the written Cantonese conventions, which constitute a widespread use within a large Cantonese linguistic community. Snow (2004, 61) highlights the growth of:

- a. the use of Cantonese in mixed texts (using Cantonese and Mandarin)
- b. the Cantonese literature

c. the different genres within it are possible to use Cantonese (generally related to oral literature).

Through times, written Cantonese appears in various scriptural objects, such as wooden fish songs, love songs from the early nineteenth century, contemporary diary-format paperback novels, Cantonese opera transcripts which are still very widespread, cartoons, street and newspaper advertisements, newspaper articles, TV dramas and digital conversation (e.g. in ICQ, emails), to name but a few (Snow 2004). This informal dimension of Cantonese echoes an important part of the Hong Kong population's feeling which is more inclined to Hongkong-ese than Chinese – A 2012 *South China Morning Post* survey concluded that only 17% of the people in Hong Kong identified themselves as Chinese. The 1970s were the starting point when a majority of Hongkongers were born in Hong Kong, contrasting the previous generation who was predominantly immigrants from the Chinese mainland or other Asian countries. In such a context, to write in Cantonese brings a young and trendy overtone as well as a political gesture against the PRC government.<sup>12</sup>

From a linguistic point of view, written Cantonese heavily borrows words from English and syntactical features from Mandarin and has minimal stability (Matthews and Yip 1994). According to Matthews and Yip:

written Cantonese is not very standardized in the sense of a widespread agreement as to which characters should be used to represent which words [although some dictionaries are published]. In fact it is not unusual for two or even more different characters to be in widespread use even for fairly common Cantonese words. (Matthews and Yip 1994, 57)

Yet, besides this outbreak of written Cantonese, a perception of Cantonese as a socially devaluated language emerges. When Cantonese is studied as a written language, Lo and Wong (1990) demonstrates that written Cantonese is perceived as a feature of popular media, and therefore, of a low standard, while written Standard Chinese is seen as a more prestigious press feature, which therefore represent a high variety); Snow assumes that Hong Kong speakers often perceive the distinction between written Cantonese and written Standard Chinese not as two different languages, but as two distinct registers (Snow 2004, 62); Bauer reports the following story when he was conducting a study on Cantonese:

I asked subjects to read aloud a story which had been written out in colloquial Cantonese with Cantonese characters used to represent Cantonese morphemes; most of the subjects performed this task without hesitation, but a few were amazed by this story and, with great seriousness, informed me that Cantonese was not a written language. (Bauer 1988, 286)

In 2012 when the corpus of the current study was collated, the same reactions from respondents were encountered. A lot of respondents in our sample noted before starting the written translation task that Cantonese is a spoken variety instead of a written language. Some of them even filled in the questionnaires in written Standard Chinese despite being constantly reminded to respond in written Cantonese. When we asked the sample if the task was difficult, some responded that the task was manageable since Cantonese is their native tongue, while some reflected that it was interesting but challenging since they rarely *wrote* Cantonese.

This under-normalisation and little stability make written Cantonese a particularly fruitful hub to test the plasticity of linguistic categories. The peripheral dimension of the language variety opens avenues for experimenting the potentiality of gender marking.

### 6.3. Data Analysis

First of all, the corpus shows important variations in translation from one respondent to another. Some of these variations embed mistakes. For example, a third-person pronoun was translated into a first-person pronoun, and some lexical items were wrongly transcribed, e.g. *fisherman* instead of *fireman*. This confirms the unfixity of written Cantonese, the absence of prescriptive tools and also the absence of norms at stake. Once mistakes are isolated, two other parts of speech are especially subject to variation, i.e. pronouns and possessive determiners, as shown in tables 6 and 7.

**Table 6: Pronoun variations**

| Nature of pronoun                   | Number of occurrences  | Number of form |
|-------------------------------------|------------------------|----------------|
| 1 <sup>st</sup> pers. sg            | 120 occurrences        | 1 form         |
| 1 <sup>st</sup> pers. pl.           | 58 occurrences         | 1 form         |
| 2 <sup>nd</sup> pers. sg. / pl.     | <b>90 occurrences</b>  | <b>5 forms</b> |
| 3 <sup>rd</sup> pers. sg. feminine  | <b>119 occurrences</b> | <b>8 forms</b> |
| 3 <sup>rd</sup> pers. sg. masculine | <b>89 occurrences</b>  | <b>6 forms</b> |
| 3 <sup>rd</sup> pers. sg. animate   | 11 occurrences         | 1 form         |
| 3 <sup>rd</sup> pers. pl.           | 57 occurrences         | 1 form         |

**Table 7: Possessive determiner variation**

| Nature of determiner            | Number of occurrences | Number of form |
|---------------------------------|-----------------------|----------------|
| 1 <sup>st</sup> pers. possessor | 30 occurrences        | 1 form         |
| 2 <sup>nd</sup> pers. possessor | 30 occurrences        | 3 forms        |

|  |                       |                |
|--|-----------------------|----------------|
| 3 <sup>rd</sup> pers. sg. F. possessed / M possessor | <b>30 occurrences</b> | <b>5 forms</b> |
| 3 <sup>rd</sup> pers. sg. M. possessed / F possessor | <b>30 occurrences</b> | <b>4 forms</b> |

It is interesting to note that variations mostly occur in gendered characters.

### 6.3.1. Pronouns

*Second-person singular pronoun, ‘You are beautiful’.* The standard second-person singular pronoun 你 *nei5* ‘you’ is used most often used in the corpus. However, among other forms, the variation 妳 *nei5*, consisting of the feminine stem 女, appears twice in the translation of the sentence. These results are in accordance with grammar description and previous literature on linguistics (Moser 1997, Matthews and Yip 1994).

In this case, the gender information is specific to Cantonese, as the source sentence in English does not offer any gender marking. The feminine version of the pronoun is therefore not hinted or constrained by the task which the respondents were expected to translate. We can thus deduce that this free variation reveals a potential grammaticalisation of the pronoun at speaker’s convenience and their interpretation of the association between gender and some adjectives, like “beautiful” in this case.

*Third-person singular pronoun, Forms of keoi5.* According to Cantonese grammar, the third-person singular pronoun 佢 *keoi5* is a generic form. In our corpus, 75% and 80% of the respondents respectively translated the third-person feminine pronoun ‘she’ and the third-person masculine pronoun ‘he’ into the same Cantonese pronoun 佢. In this case, the translational process showed eradication of gender information.

It is noteworthy that gender shunning is slightly more frequent for ‘he’ (80%) than for ‘she’ (75%). This implies that the generic facet of a pronoun is more easily associated with the masculine than the feminine, which needs more effort to be specified. These observations meet Michard’s hypothesis about French that the masculine semantically encodes primarily /human/ and secondarily /male/, while the feminine encodes primarily /female/, and secondarily /human/ (1999). Moser (1997) draws the same conclusion for Mandarin. After having studied a corpus of jokes, he demonstrates that *rén* ‘person’ is a covert masculine lexical item: the unmarked generic takes the role of a generic, while the feminine, as marked, is restricted to women; and “‘persons’ are by large men and thus this is the female that is perceived as requiring a special marking” (Moser 1997, 13).

20-25% of the translation of English pronouns displays respondents’ individual strategies to maintain gender features. Such realisation illustrates the intrinsic possibility of adding gender to Cantonese pronouns which are

normally not inflectional. We are witnessing here the plasticity of the pronominal category through potential realisation of gender in grammar. It has to be kept in mind, nevertheless, that such realisation occurred because of certain translation constraints and the unfixedness of written Cantonese. What, hence, are the gendering strategies employed by the respondents to intentionally exhibit gender information?

Four types can be found in total:

(1) Some speakers removed the ‘person’ radical 亻 from the male/generic pronoun 佢 and affixed the feminine radical 女 to the phonetic component 𠬞 *geoi6*, generating a new feminine character: 姪 *geoi6*. Interestingly enough, the character 姪, which is the name of a mountain in Mainland China, was transformed by some respondents for a new purpose: to explicitly indicate the feminine gender of a third-person singular noun.

(2) Some respondents used Chinese (Mandarin) third-person pronouns: 她 *taa1* [feminine] and 他 *taa1* [masculine] to translate ‘she’ and ‘he’ respectively.

(3) Other respondents chose lexicalised forms to translate the English pronouns. Examples are as follows:

- |                              |                      |
|------------------------------|----------------------|
| (5) 個女仔 <i>go-neoizai</i>    | ‘this girl’          |
| 女仔 <i>neoizai</i>            | ‘girl’               |
| 佢女朋友 <i>keoi-neoipangjau</i> | ‘his/her girlfriend’ |
| 男仔 <i>naamzai</i>            | ‘boy’                |
| 男嘅 <i>naamge</i>             | ‘this man’s’         |

The data provided above indicate that, although the process of genderisation is triggered by the task, different possibilities regarding the marking and unmarking of gender and the ways of marking it emerge (that is, borrowing to Chinese, lexicalisation of feminisation by neologism). The results of the translation of second- and third-person pronouns show that written Cantonese pronominal system potentially contains gender marks. The heterogeneity of strategies deployed by the respondents makes gender a prospective classifier.

*Anaphora: Tony, Jenny and the dog*, We could see in the above examples that some pronouns are potentially “genderable” at the respondents’ convenience. What actually happens within an ambiguous anaphoric context, where gender information is relevant regarding the syntactical organisation of English, can be interpreted through the translation of the following sentence:

*Tony and Jenny are going to Macau. He wants to try the egg tart there, while she wants to do some sightseeing.*

Twelve respondents out of thirty refuse the available anaphoric options but rather chose a nominal resumption. This is the sign of reluctance to mark gender on the pronoun while maintaining gender information. The most common pattern of translation of the second part of the above sentence is equivalent in English to: 'Tony wants to... while Jenny wants to' which features the repetition of the two proper nouns instead of replacing them with pronouns.

Seven respondents chose to ignore the gender information of the source sentence by using the same pronoun 佢 for the two anaphora. An English word-for-word equivalent of the Cantonese translation would be 'They wants to ... while they wants to'.<sup>13</sup>

Nine in the sample alternately used nominal and pronominal forms in order to disambiguate the anaphoric distribution, rendering some versions of translation equivalent to 'Tony wants to ... while she/ they wants to' or 'He/they wants to... while Jenny wants to'. In this case, the respondents avoided gender grammaticalisation (pronominalisation) while maintaining the semantic gender distinction.

Only one respondent highlighted the gender of 'Tony' and 'Jenny' by means of the gendered pronouns 佢 and 她 respectively.

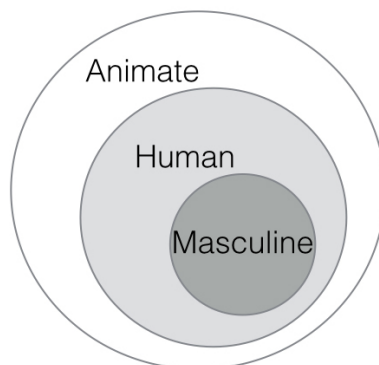
A majority of respondents adopted thus strategies to maintain the gender information, without resorting to anaphora. These strategies are lexicalisation or alternation between lexical and pronominal items. In this case, pronominal gendering is marginalised.

In this sentence:

*The dog is tired now because it played all day.*

The anaphora concerns a non-human animate referent. Eleven respondents used 佢 as the translation of 'it'. This result corroborates the findings of Chan (2001) who discovers that the pronoun 佢 encodes primarily an animate feature rather than a human or gender feature. With reference to Michard (1999) and Chan (2001), it therefore seems that the pronoun 佢 bears a semantic organisation relying on a primary vs. secondary meaning while having a central vs. peripheral meaning. Figure 1 represents the semantic repartition, with the biggest circle representing the primary meaning, and the darkest circle representing the most central meaning. Primary and central meanings are amalgamated through the generic dimension.

**Fig. 1: Semantic organisation of the pronoun 佢**



### 6.3.2. Genitive

The gender allocation in genitive marking was also investigated through the following sentence:

*She is his friend.*

Twenty-six out of thirty respondents (87%) erased the gender information in the translational process and made use of the “generic” implication of 佢 for ‘she’, and 佢, 佢個 *keoi-go* or 佢嘅 *keoi-ge* for ‘his’ (個 *go3* and 嘅 *ge3* encode the genitive), where no mark of gender appeared.

Only one speaker marked the masculine in translating the genitive ‘his friend’ by 呢個男仔 *neigo-naamzai*, within which the two latter Chinese characters 男仔 mean ‘boy’ (a word-for-word transliteration equivalent would be: ‘they is the boy’s friend’).

In three translated sentences, the feminine feature of the subject ‘she’ was maintained using either the invented pronoun 姪 or the written Standard Chinese feminine pronoun 她, or:

- |     |                  |             |            |
|-----|------------------|-------------|------------|
| (6) | 個                | 女           | 仔          |
|     | <i>go-</i>       | <i>neoi</i> | <i>zai</i> |
|     | [genitive]       | [female]    | [boy]      |
|     | ‘the boy’s girl’ |             |            |

These results confirm the tendency of Cantonese to shun gender information in grammatical morphemes such as subjects and genitive pronouns. It appears that in rare cases where gender was maintained, it was done on the nominative pronoun rather than on the genitive.

In the above example, we again witness a phenomenon that, whichever genderisation strategy was adopted, feminine identities are more frequently specified than the masculine.

*Her boyfriend is an American.*

Similarly, when dealing with ‘her’, 96% of the translation of the above sentence displays a neutral third-person mark 佢, with only one occurrence of an explicit feminine mark using written Standard Chinese 她. It should be



noted that this sole occurrence does not come from the same respondent who also replaced ‘she’ with the same Standard Chinese feminine pronoun in the previous sentence.

Regarding the translation of *boyfriend*, dissimilar to the previous example, twenty-five out of thirty respondents provided masculine forms in three variations: 男朋友 *naam-pangjau*, 男友 *naam-jau* or 條仔 *tiu-zai*, of which 男 and 仔 encode ‘male’).

Therefore, in English, such a sentence displays two gender marks (*her* [feminine], *boyfriend* [masculine]), while in Cantonese only one gender mark can be seen (男朋友, ‘boyfriend’ [masculine]). The elimination of one gender mark could be explained by structural motives together with the social ones. On the formal side, this example and the previous one suggest a tendency of Cantonese speakers to mark gender on lexical materials rather than on grammatical materials such as pronouns; and when needed, to mark gender on nominative pronouns rather than on genitive pronouns. From a social point of view, a heteronormative context might lead one to think that the translation of *boyfriend* as 男朋友 implies a feminine possessor.

We assume here that the degree of linguistic embedment of gender is not linked to the degree of gender strength in a given society. Gender lexicalisation is rather one of the possible modalities of doing/saying gender. However, as it happens, this grammatical unmarking of gender could represent avoidance of gender grammaticalisation as well as speakers’ belief in heterosexual social norms. The corpus, consequently, could be enriched by adding sentences such as *His boyfriend is an American* for translation to see what happens.

### 6.3.3. Nouns

Respondents’ translation shows a huge variety of the interpretation of nouns which possess no gender attachment. For example, there are 10 different translated versions of *brother*. This is due to a more sophisticated lexicon for kinship in Cantonese than in English, commanding the respondents to add information not specified in the English equivalent. Besides, similar to English, nouns of relationship such as *father* and *brother* do not carry grammatical mark of gender. Gender is lexically encoded at a primary level.

*American*. 97% of its translation displays a neutral form 美國人 *meigwok-jan* with the use of a generic noun 人 *jan4* referring to human, while only one respondent attached an aforementioned gendered noun 仔 to 美國 *meigwok* ‘America’, supplying 美國仔 *meigwok-zai* ‘American boy’ as the translation. Unlike the genitive cases studied above, the masculine feature here does not bring any semantic presumption related to another (feminine) referent, which indicates that the use of masculine does not serve as a disambiguation. In the sentence *Her boyfriend is an American*, the

masculine feature of the referent has already been displayed in the translation of *boyfriend*, i.e. 男朋友. It can be seen here that the choice of this respondent has nothing to do grammaticalisation or disambiguation.

*Guy*. Conversely, it is intriguing to note that Cantonese strongly unmarks gender in the following sentence: *He is a very nice guy*. While there are two gender marks of a common referent in English, namely *he* and *guy*, twenty out of thirty respondents (67%) gave an ungendered translation in Cantonese. *He* is predominantly translated into 佢 (generic/masculine) and *guy* into 人 ‘person/someone’. There are only seven occurrences of the masculine form 男仔 for the translation of *guy*. This is again in agreement with the analysis that Michard coins for French (1999) that masculine terms primarily signify a generic feature and secondarily a male feature, whereas feminine terms first specify a gender information (female), and only means ‘human’ secondarily.

*Actress*. It is observed that the case of *actress* is in opposition to other examples discussed above. In this English sentence with one gender attachment: *This actress speaks loudly*, nine respondents, which is close to one-third of the corpus, hid the gender information by giving 演員 *jinjyun* or 藝人 *ngaijan*, both of gender-neutrally meaning ‘performer/artiste’, as the translation, while 70% of them preferred affixing 女 to the neutral referents of ‘actress’: 女主角 *neoi-zyugok* ‘female protagonist’ and 女藝員 *neoi-ngaijyun* ‘female artiste’. These opposite choices which concern feminine referents again confirm Michard’s analysis: it is more difficult to erase gender information when referring to a feminine term, while it is easier to un-specify it for a male referent, where both masculine and generic forms can be used.

*Friend, writers, etc.*

Here, the contrastive analysis does not show any gender shift. Paralleling English, Cantonese does not add specific gender to these terms, regardless of which they refer to and whether there is gender information in other parts of the same sentences.

*Non-humans terms*

Unsurprisingly, non-human referred terms like *tea, book, dog, Coca-Cola, philosophy, Cantonese, Macau* are totally devoid of gender, and these responses are again in favour of the lexical classifier hypothesis discussed in section 2.

#### 6.3.4. Articles/Quantifiers

English-Cantonese translation requires switching English articles to Cantonese quantifiers, which are a kind of classifiers. The corpus shows that quantifiers in Cantonese are distributed according to a human/non-human and animate/inanimate distinction. In the survey, the English determiner *this* was translated into 個 *go3* when respondents recognised human-referred terms in sentences, while the same determiner was translated into 本 *bun2* and other matching quantifiers for non-human heads. Similarly, the definite article *the* in English was transposed to 隻 *zek3* for the animate, but nil ( $\emptyset$ ) for the inanimate. There is therefore a grammatical encoding of the (in)animate/(non)human feature. The distribution of the data is summed up in table 8:

**Table 8: Quantifiers semantic features**

|             | Human | Non-human<br>animate | Inanimate   |
|-------------|-------|----------------------|-------------|
| <i>this</i> | 個     |                      | 本           |
| <i>the</i>  |       | 隻                    | $\emptyset$ |
| <i>a</i>    | 個     |                      |             |

A more advanced analysis has to be done in order to provide more precise and accurate categorisation of semantic embedment in Cantonese morphosyntax. Is the animate feature primary semantic feature, as Chan (2001) postulates, or is it the human one which is the primary one? This is still worth further exploration, since the results would serve as a useful tool to understand against which semantic-syntactical backdrop gender spreads in Cantonese.

#### 6.4. Analysis sum-up

Though the huge variation of gender marking in written Cantonese and the exploratory dimension of our corpus make us cautious at times when summing up our findings, valuable implications are still retrievable. We postulated the inter-lingual space as an observation room for gender form shifts; we formulated a hypothesis that gender transpositions would enlighten minor phenomena and revealed categorial plasticity. Such plasticity constitutes precisely the shape of the gender form.

The English-Cantonese translation corpus exhibits the tendency for the unmarking of gender in Cantonese. That is, in comparison with English, Cantonese provides less gender information – gender information in Cantonese is, in most cases, lost or at least covert. This first finding allows us to conclude that in the two target languages which are commonly known

to be genderless, different degrees of gender embedding can still be observed. But as soon as we look at the detail of this loss at an inter-level scale, we are able to notice that this “gender loss” concerns mostly the morpho-syntactical level but is often balanced by gendering at the lexical level. The gender form is in fact translated by means of a move through and across linguistic levels. In order to transfer gender from pronouns to lexicons, speakers of Cantonese employ different strategies. For example, while refusing anaphoric constructions, they resort to nominal resumption lexicalisation (see the *Tony and Jenny* example) and lexical combinations (see the *Actress* example); they sometimes make use of Standard Chinese or even draw on neologisms. Exceptionally, these strategies also allow the addition of gender information to words/sentences which are not present in English (see the *American* example). The implication here is that gender in Cantonese is rather encoded at a compositional level rather than a derivational/inflectional level (morphosyntactical). The form of gender is thus transposed through the different linguistic levels in the Cantonese language.

Some peripheral marking of gender adopted by respondents reveal that this tendency is not due to a morphological blockage of Cantonese. Indeed, pronouns are occasionally gendered by Cantonese speakers. If one could expect such a phenomenon for the second-person pronoun, it is more surprising for the third-person pronouns. A morphological gender is then linguistically possible, although not structuralised. This linguistic potentiality of gender is to be replaced in the context of the unfixedness of written Cantonese due to its emerging feature. This periphery also reveals the strong interplay between syntax and the semantic cut-up of reality. The intrinsic tension between form and meaning reminds us that the study of gender in a language cannot afford to take for granted what is structural in a language and what is not. Linguistic structure has to be historicised, and this point has been illustrated through the emergence of a gender contrast for the second-person pronoun which is linked to certain language policy agenda.

Finally, the corpus analysis has shown that both elimination and display of gender are not only level-sensitive, but above all gender-sensitive, according to the semantic cut-up of Cantonese. If the animate/inanimate opposition seems the primary one, it appears that the animate category is centralised around the masculine, assigning the feminine to the periphery. This is why the feminine needs to be specified more often than the masculine, and such observation proves that the amalgam of masculine and generic features also occurs in languages, within which gender is slightly grammaticalised.

This analysis is to be supplemented by a broader corpus allowing a more thorough textual analysis. Indeed, by working on more extended resources, we could work, on the one hand, on the generic/masculine relationship and explore, on the other hand, if the emergence of feminine pronoun shifts the grammatical meaning of the generic pronoun from neutral to masculine. If it

happens to be the case, a feminist perspective of such women's visibilisation has to pay attention to avoid the freezing of this new categorisation in a renewed essentialism. Additionally, a textual analysis would offer an opportunity to discuss at a rhetorical level the implicit and explicit features of gender (see the discussion about the "*her boyfriend*" sentence).

## 7. Conclusion

It has been proposed in this chapter to examine gender in Cantonese at the crossroads of the linguistic tradition focusing on the syntactical features of gender and the gender studies, focusing on the semantic dimensions of gender. We suggested a hermeneutical approach, within which the linguistic structure involved is not a mere formal system but a process of solidification and institutionalisation of the gender categories in languages. This approach, combining corpus observations and theoretical discussion, led us to portray gender in a multidimensional way:

- a. scalarity of gender grammaticalisation,
- b. semiotic dimension as the primary dimension of gender,
- c. structuralisation as a solidification of the gender categorisation process,
- d. definition of gender as a form, that is a transposable relationship, and
- e. *nannü youbie*, that is the dichotomous distinction as a power relationship founding the gender categories, as proposed by Zhen.

Based on these features, we outlined the need of identifying the gender form as an inter-linguistic levels continuum across languages, leading us to a contrastive approach between English and Cantonese.

Such a semiotic definition of gender utilises language structure as a site where power-relationships are deployed. This statement implies that attention should be paid to structuralisation rather than the structure itself. For this purpose, we ought to take a historicised vision of linguistic structures. The first part of the study reveals to this regard that English and Cantonese present reverse tendencies regarding gender grammaticalisation, with a late and light genderisation of grammar in Cantonese (especially pronominal system), and a gender loss in English syntax.

This structuralisation process could also be observed in the corpus. If English and Cantonese present different degrees of gender grammaticalisation, the transposition of gender forms actually appears by means of a move through different linguistic levels, from morphosyntax to lexicon. This move was apparently attributable to the structural and normative plasticity of the gender categories in written Cantonese, a particularly non-standardised variety. The data analysis of the translation survey also reveals that the gender elimination or display is gender-sensitive. The animate category is centralised around the masculine,

assigning the feminine to the periphery. In this way, feminine marks have to be more explicitly specified than masculine ones, which can conveniently be expressed through the generic form. This proves that the amalgam of masculine and generic features also occurs in language whose gender is lightly grammaticalised. The strategies at stake to genericise the feminine or specify the masculine in Cantonese, though beyond the scope of this paper, would provide a fruitful insight into the linguistic structure negotiation for political purposes.

### Acknowledgement

This work has been carried out thanks to the support of VariAMU, an A\*MIDEX project (n° ANR-11-IDEX-0001-02) funded by the “Investissements d’Avenir” French Government program, managed by the French National Research Agency (ANR).

### References

- Angenot, Marc. 1982. *La parole pamphlétaire : contribution à la typologie des discours modernes*. Paris: Payot.
- Baron, Dennis. 1986. *Grammar and Gender*. Yale: Yale University Press.
- Bauer, Robert S. 1988. “Written Cantonese in Hong Kong.” *Cahiers de linguistique - Asie Orientale* 17(2): 245-93.
- Blache, Philippe, Henning Christiansen, Verónica Dahl, Denys Duchier, and Jørgen Villadsen. 2014. *Constraints and Language*. Cambridge : Cambridge Scholars Publishing.
- Bourgeois, Pénélope. 1978. “Simplification des caractères chinois : ‘Ebauche’ du 2ème projet de simplification.” *Cahiers de linguistique - Asie Orientale* 4(1): 89-103.
- Burr, Elisabeth. 2012. “Planification linguistique et féminisation.” In *Intersexion. Langues romanes, langue et genre*, ed. by Fabienne Baider, and Daniel Elmiger, 29-40. Munich: Lincom.
- Butler, Judith. 1990. *Gender Trouble*. New York: Routledge.
- Cameron, Deborah. 1985. “What Has Gender Got to Do with Sex?” *Language and Communication* 5(1): 19-27.
- . 1990. *The Feminist critique of language : A Reader*. London; New York: Routledge.
- . 1992. *Feminism and linguistic theory*. New York: St. Martin’s Press.

- Census and Statistics Department of the Hong-Kong Government. 2012. <http://www.censtatd.gov.hk/home/>
- Chan, Marjorie K. M. 1999. "Sentence-Final Particles in Cantonese: A Gender-Linked Survey and Study." In *Eleventh North American Conference on Chinese Linguistics*, ed. by Baozhang He, and Wenzhe Hu, 87-101.
- Chan, Thomas. 2001. *Orthographic change: Yue (Cantonese) Chinese dialect characters in the nineteenth and twentieth centuries*. PhD dissertation. Ohio: Ohio State University.
- Cheng, Anne. 1997. *Histoire de la pensée chinoise*. Paris: Le Seuil.
- Chevalier, Yannick. 2013. "Approches linguistiques du genre (gender)." In *Apéro'Confs*. Lyon. <http://cle.ens-lyon.fr/plurilingues/approches-linguistiques-du-genre-gender--214114.kjsp>.
- Corbett, Greville G. 1991. *Gender*. Cambridge: Cambridge University Press.
- . 2013. "Sex-based and Non-sex-based Gender Systems." In *The World Atlas of Language Structures Online*, ed. by Matthew S. Dryer, and Martin Haspelmath. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/31>
- Curzan, Anne. 2003. *Gender Shifts in the History of English*. Cambridge; New York: Cambridge University Press.
- Egerod, Søren. 1982. "Sino-Tibetan Languages." *Encyclopaedia Britannica*, 796-806.
- Erbaugh, Mary. 2002. "Classifiers are for specification: complementary functions for sortal and general classifiers in Cantonese and Mandarin." *Cahiers de linguistique - Asie Orientale* 31(1): 33-69.
- Fung, Anthony. 2006. "Gender and Advertising: The Promotional Culture of Whitening and Slimming." In *Advertising in Hong Kong Society*, ed. by Kara Chan, 171-82. Hong Kong: The Chinese University of Hong Kong Press.
- Grinevald, Colette. 1999. "Typologie des systèmes de classification nominale." *Faits de langues* 14: 101-22.
- Hellinger, Marlis, and Hadumod Bussmann. 2001. *Gender across Languages* vol. 1. Amsterdam: John Benjamins.
- . 2002. *Gender across Languages* vol. 2. Amsterdam: John Benjamins.
- . 2003. *Gender across Languages* vol. 3. Amsterdam: John Benjamins.
- Huddleston, Rodney, and Geoffrey Pullum. 2008. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

- Jakobson, Roman. 1971. *Selected Writings, World and Language* vol. 2, ed. by Stephen Rudy. The Hague; Paris: Mouton de Gruyter.
- Kane, Daniel. 2006. *The Chinese Language: Its History and Current Usage*. North Clarendon: Tuttle Publishing.
- Kang, Agnes. 2008. "At the intersection of elitism and gender in Hong Kong: advertisements of luxury residences." In *Igala 5 Proceedings*, 97-106. Wellington: Victoria University of Wellington.
- Kang, Agnes, and Katherin Chen. 2012. "Stancetaking and the Making of the (Hong) Kong girl." In *Sociolinguistics Symposium 19*. Berlin.
- Kibort, Anna, and Greville G. Corbett. 2008. "Gender." *Grammatical Features*. <http://www.grammaticalfeatures.net/features/gender.html>.
- Ku, Agnes S. 2001. "Hegemonic construction, negotiation and displacement: The struggle over right of abode in Hong Kong." *International Journal of Cultural Studies* 4(3): 259-78.
- Lado, Robert. 1957. *Linguistics Across Cultures*. Ann Arbor: University of Michigan Press.
- Law, K.W.K., and A.H.N. Chan. 2004. "Gender role stereotyping in Hong Kong's primary school Chinese language subject textbook." *Asian Journal of Women's Studies* 10(1): 49-69.
- Lee, Jackie F.K., and Peter Collins. 2008. "Gender Voices in Hong Kong English Textbooks—Some Past and Current Practices." *Sex Roles* 59: 127-37.
- Lee, Keekok. 2008. *Warp and Weft. Chinese Language and Culture*. New York: Eloquent Books.
- Lewis, Paul M., Gary F. Simons, and Charles D. Fennig. 2014. *Ethnologue: Languages of the World, Seventeenth edition*. Dallas: SIL International. <http://www.ethnologue.com>
- Liu, Lydia H., Dorothy Ko, and Rebecca E. Karl. 2013. *The Birth Of Chinese Feminism. Essential Texts in Transnational Theory*. New York: Columbia University Press.
- Lo, Terence, and Colleen Wong. 1990. "Polyglossia in the 'printed cantonese' mass media in Hong Kong." *Journal of Asia Pacific Communication* 1(1): 27-43.
- Martinet, André. 1970. *Éléments de linguistique générale*. Paris: Armand Colin.
- . 1999. "Genre et sexe." *La Linguistique* 35(2): 5-9.
- Matthews, Stephen, and Virginia Yip. 1994. *Cantonese: a comprehensive grammar*. London: Routledge.



- Matthiessen, Christian M.I.M., and Michael A.K. Halliday. 2009. *Systemic Functional Grammar: a First Step Into the Theory*. Beijing: Higher Education Press.
- Michard, Claire. 1999. "Humain/femelle : deux poids deux mesures dans la catégorisation de sexe en français." *Nouvelles Questions Féministes* 20(1): 53-95.
- Moser, David. 1997. "Covert Sexism in Mandarin Chinese." *Sino-Platonic Papers* 74: 1-23.
- Motschenbacher, Heiko. 2010. *Language, Gender and Sexual Identity: Poststructuralist Perspectives*. Amsterdam; Philadelphia: John Benjamins.
- Pan, Wenguo, and Wai Mun Tham. 2007. *Contrastive Linguistics. History, Philosophy and Methodology*. London; New York: Bloomsbury.
- Pauwels, Anne. 2010. "Socially Motivated Language Reform in a Global Lingua Franca: The Case of Gender Reform in English." In *Language in Its Socio-Cultural Context, New Explorations in Gendered, Global and Media Uses*, ed. by Markus Bieswanger, Heiko Motschenbacher, and Susanne Mühleisen, 21-33. Frankfurt am Main: Peter Lang.
- Peyraube, Alain. 1999. "Historical change in Chinese Grammar." *Cahiers de linguistique - Asie Orientale* 28(2): 177-226.
- . 2001. "Syntaxe et discours dans la langue chinoise: le mythe de l'occidentalisation." *Mots. Les langages du politiques* 66: 19-32.
- Poon, Anita Y.K. 2004. "Language policy of Hong Kong: Its impact on language education and language use in post-handover Hong Kong." *Journal of Taiwan Normal University: Humanities & Social Sciences* 49(1): 53-74.
- Rastier, François. 2001. "Indécidable hypallage." *La Langue française* 129 (1): 111-27.
- Romaine, Suzanne. 2001. "A Corpus-Based View of Gender in British and American English." In *Gender across Languages* vol. 1, ed. by Marlin Hellinger, and Hadumod Bussman, 153-75. Amsterdam: John Benjamins.
- Schnurr, Stephanie. 2008. "'I'm such a small potato' gender leadership discourse in Hong Kong workplace." In *Igala 5 Proceedings*, 133-46. Wellington: Victoria University of Wellington.
- . 2010. "'Decision Made – Let's Move on' Negotiating Gender and Professional Identity in Hong Kong Workplaces." In *Language in Its Socio-Cultural Context, New Explorations in Gendered, Global*

- and Media Uses*, ed. by Markus Bieswanger, Heiko Motschenbacher, and Susanne Mühleisen, 111-35. Frankfurt am Main: Peter Lang.
- Scott, Joan W. 1986. "Gender: A Useful Category of Historical Analysis." *The American Historical Review* 91(5): 1053-1075.
- Silverstein, Michael. 1985. "Language and the culture of gender: at the intersection of structure, usage and ideology." In *Semiotic Mediation*, ed. by E. Mertz, and R.J. Parmentier, 219-59. Orlando: Academic Press.
- Snow, Don. 2004. *Cantonese as Written Language*. Hong Kong: Hong Kong University Press.
- Violi, Patrizia. 1987. "Les origines du genre grammatical." *Langages* 85: 15-34.
- Voisin, Sylvie. 2014. "Classes nominales et complexité du système." In *Séminaire Langues en contact et Typologie*. Aix-en-Provence.
- West, Candace, and Don Zimmerman. 1987. "Doing Gender." *Gender and Society* 1(2): 125-51.
- Whorf, Benjamin L. 1937. "Grammatical Categories." In *Language, Thought, and Reality. Selected Writings of Benjamin Lee Whorf*, ed. by John B. Carroll, 87-101. Cambridge: The MIT Press.
- . 1940. "Science and Linguistics." In *Language, Thought, and Reality. Selected Writings of Benjamin Lee Whorf*, ed. by John B. Carroll, 207-19. Cambridge: The MIT Press.
- Wong, Andrew D. 2005. "The reappropriation of tongzhi." *Language in Society* 34(5): 763-793.
- Yang, Chi Cheung Ruby. 2011. "Gender representation in a Hong Kong primary English textbook series: The relationship between language planning and social policy." *Current Issues in Language Planning* 12(1): 77-88.
- Yang, Zijian. 2007. "Fresh Developments in Contrastive Linguistics." In *Contrastive Linguistics*, ed. by Pan Wenguo and, Tham Wai Mun, 1-19. London; New York: Bloomsbury.

---

<sup>1</sup> We warmly thank Karim Hammou for his inspiring discussion and wise advice. We'd also like to acknowledge Virginia Yip for her useful reminders regarding the sensitive dimensions of gender in Cantonese. All imprecisions or mistakes remain ours.

<sup>2</sup> "Deux orientations s'offrent à nous si nous voulons étudier le *genre* dans ses rapports avec le *genre* : soit envisager ce que les études de genre doivent à la grammaire, soit étudier ce que la grammaire doit aux études de genre."

---

<sup>3</sup> *Chauffeuse* is also in use with a human reference acceptation ('female driver'), where the feminine mark is semantically motivated. We consider this acceptation as a synonym and do not address the term in its human-referred meaning.

<sup>4</sup> Our italics

<sup>5</sup> We use here functionalist in the sense of Martinet (1970; 1999), rather than of Mathiessen and Halliday (2009).

<sup>6</sup> We have to note here that such a formalist understanding of the linguistic structure is nowadays more and more challenged, and formalists pay an increasing attention to uses and variation phenomena.

<sup>7</sup> Cantonese pinyin is given according to the Chinese Character Database of the Chinese University of Hong Kong <http://humanum.arts.cuhk.edu.hk/Lexis/lexi-can/>

<sup>8</sup> Yip suggests that such constructions may have appeared under West language influence (personal communication 2013). See below for a discussion of this point.

<sup>9</sup> Cantonese and Mandarin are two non-mutually understandable languages. However, they share the same scriptural system, respectively in its traditional or simplified version.

<sup>10</sup> See He-Yin Zhen's texts, edited by Karl, Ko and Liu (2013) for a precise description of the Chinese reception of Western narratives about women's liberation.

<sup>11</sup> We rely heavily in this section on the works of Snow (2004) and Bauer (1988), which are the major (and almost exclusive) studies available about written Cantonese. We have to, however, express reservations about Snow's reference. Although it is indisputably a pioneer and ground-breaking work providing a relevant sociolinguistic overview on written Cantonese, it misses a solid analysis based on linguistic facts. We thus use this reference for its general and factual information.

<sup>12</sup> This statement seems to have been socially actualised in view of the recent events happening in Hong Kong in Autumn 2014 as a reaction against the potential intrusion of PRC in Hong Kong's political reform. It is, however, too early to make any claim regarding Hong Kong's identity shift and its influences on the use of written Cantonese in the city.

<sup>13</sup> We use here the singular implication of '*they*' to render the generic dimension.