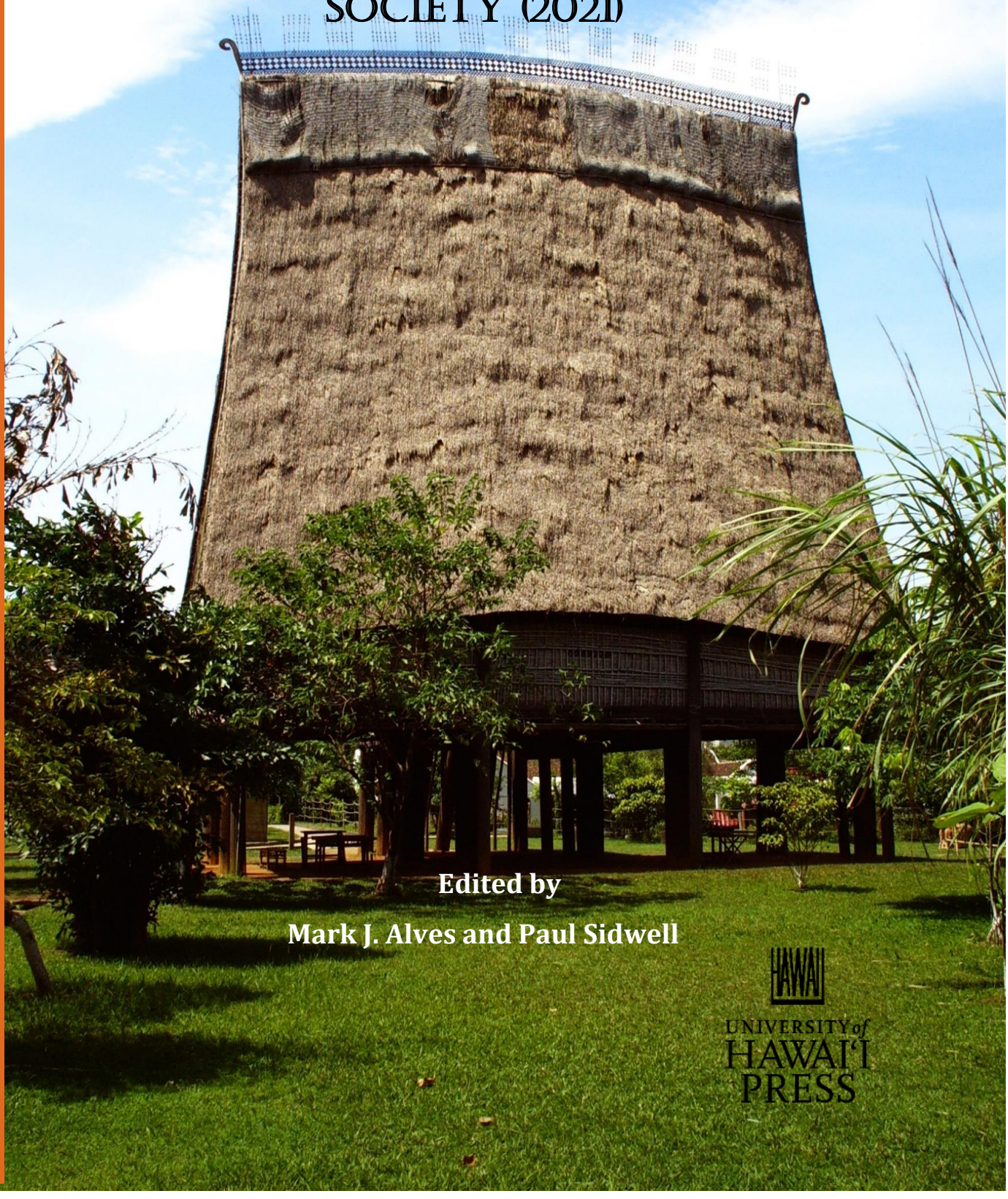
JSEALS Special Publication No. 8

# PAPERS FROM THE 30TH MEETING OF THE SOUTHEAST ASIAN LINGUISTICS SOCIETY (2021)



Edited by

Mark J. Alves and Paul Sidwell

UNIVERSITY of HAWAI'I PRESS

OPEN ACCESS – Semiannual with periodic special publications
E-ISSN: 1836-6821
http://hdl.handle.net/10524/52498



Cover photo adapted from the cover of the JSEALS journal provided by Paul Sidwell.

# JSEALS
Journal of the Southeast Asian Linguistics Society



The Journal of the Southeast Asian Linguistics Society publishes articles on a wide range of linguistic topics of the languages and language families of Southeast Asia and surrounding areas. JSEALS has been hosted by the UH Press since the beginning of 2017.

# CONTENTS

**Phonetics and/or Phonology**

# NONCANONICAL PRONOUNS IN VIETNAMESE AND CHINESE

Chenchen SONG
*Zhejiang University*
*cjs021@zju.edu.cn*

Li NGUYEN
*University of Cambridge*
*nhbn2@cl.cam.ac.uk*

**Abstract**
In this paper, we study a new type of pronominal item emerging on the Internet in Vietnamese and Chinese. First, we demonstrate that pronominal items of this new type, which we dub "noncanonical", are a separate category from both textbook default pronouns and imposters (Collins and Postal 2012). Then, we illustrate their real-life usage in detail. Our investigation shows that noncanonical pronouns in the two Asian languages are similar not only in syntactic behavior but also in lexical sources, based on which we propose three subtypes for them. Finally, we account for the half-grammatical-half-lexical status of noncanonical pronouns in the theory of generalized root syntax (Song 2019), a recent version of syntactic root theory. We also suggest a link between the propensity for noncanonical grammatical elements and high analyticity.

**Keywords:** Vietnamese, Chinese, pronoun, Internet, root syntax, analyticity
**ISO 639-3 codes:** vie, zho

## 1. Introduction

### *1.1 Pronominal items in previous research*

Previous research has documented two types of personal pronominal items in human language. The first type is the default personal pronoun (henceforth default pronoun)[1]—namely, the kind of pronoun typically seen in language textbooks and reference grammars, as exemplified in (1).

(1) a. English: I, you, he, she, it…
b. German: *ich, du, er, sie, es…*
c. Vietnamese: *tôi, bạn* (lit. 'friend')*, anh ta, cô ta…*
d. Mandarin: *wǒ, nǐ, tā* ('he/she/it')...

Such pronouns are not necessarily the most often used, especially in languages like Vietnamese, but they are deemed textbook standards and are also the most widely studied type of pronominal item in linguistics. Textbook pronouns can be viewed as exponents of formal features, especially person, number, and gender (i.e., phi) features and occasionally also honorific features, as exemplified in (2).

(2) a. **I** am a student. (I = [1SG])
b. *Möchten* ***Sie*** *etwas* *zu* *trinken?* [German]
would.like.3PL you.HON something to drink
"Would you like something to drink?" (*Sie* = [2SG/PL, HON])

[1] Since we are only concerned with personal pronouns in this study, for simplicity's sake we use "pronoun" to mean "personal pronoun" throughout. By "default" we mean to contrast the kind of pronominal item in (1) with the imposters and especially the noncanonical pronouns to be introduced below.

In (2a), the English pronoun *I* is the exponent of the featural specification [1SG]. In (2b), the German pronoun *Sie* is the exponent of [2SG/PL, HON]. A hallmark of textbook pronouns is that their grammatical behavior can be explained solely by their formal features.

The second type of pronominal item documented in previous research is the "imposter" (Collins and Postal 2012). An imposter looks like an ordinary referring expression (R-expression) and is subject to third-person agreement in languages where syntactic agreement is required, but it semantically refers to the speaker or the addressee instead of a real third person. See (3) for an illustration.

(3) a. **Daddy** (= I) is going to get you an ice-cream cone.
b. Is **the general** (= you) going to dine in his suite. (Collins & Postal 2012:1–3)

Imposters carry more pragmatic content than textbook pronouns. Thus, *daddy* in (3a) sounds more affectionate than *I*, and *the general* in (3b) sounds more formal than *you*. Descriptively we can say that imposters are R-expressions employed to refer to the speaker/addressee. There are many more imposters in English, such as *yours truly* (= I), *this reviewer* (= I), *Madame* (= you), *sweetie* (= you), etc. Imposters are also widely attested in other languages (see the articles in Collins 2014). In fact, they are used so frequently and naturally in some languages, including Vietnamese, that they may be viewed as the de facto default pronominal items there, taking over the role of textbook standard pronouns. Imposters have not taken over the role of textbook pronouns in Chinese but are also highly common. See (4) to (6) for a comparative illustration of Vietnamese and Chinese[2] imposters, which fall in three subtypes: kin terms, career titles, and personal names.

(4) a. *Hôm-nay* ***mẹ*** *sẽ* *nghỉ* *làm.* [Viet.]
today mom.1SG FUT stop work
"Mom (= I) is having a day off today." (a mom talking to her child)
b. *Jīntiān* ***māma*** *zuò* *yú-tāng.* [Mandarin]
today mom.1SG make fish-soup
"Today mom (= I) will cook fish soup." (a mom talking to her child)

(5) a. ***Thầy*** *về-hưu* *rồi.* [Viet.]
teacher.1SG retire PRF
"Teacher (= I) has retired." (a teacher talking to their student)
b. ***Lǎoshī*** *yě* *bù* *zhīdào* *dá'àn.* [Mandarin]
teacher.1SG also not know answer
"Teacher (= I) don't know the answer either." (a teacher talking to their student)

(6) a. ***Linh*** *hiểu* *chưa?* [Viet.]
Linh.2SG understand IMPERF
"Has Linh (= you) understood yet?" (the addressee's name is Linh)
b. ***Língling*** *bù* *chī* *qiǎokèlì* *ma?* [Mandarin]
Lingling.2SG not eat chocolate Q
"Lingling (= you) don't eat chocolate?" (the addressee's name is Lingling)

Although previous studies on Vietnamese rarely (if ever) use the term "imposter", that the boldfaced items in (4a) to (6a) qualify as imposters is clear: they are R-expressions employed to refer to the speaker/addressee.

[2] Unless otherwise specified, all our Chinese data are from Standard Mandarin. We also use the term Common Mandarin (i.e., the Mandarin variety commonly spoken in daily life) when discussing Internet language phenomena that have not yet been officially recognized as part of Standard Mandarin.

Overall, (4) to (6) reveal that imposters are used a lot more freely in Vietnamese[3] and Chinese than in English and that as such they are not a crosslinguistically homogeneous phenomenon.[4] A most noticeable point of variation concerns agreement patterns. Since Vietnamese and Chinese both lack formal agreement, their imposters do not manifest the kind of mismatch between grammatical and notional person discussed at length in Collins and Postal (2012). Note that while the English translations all manifest third-person agreement (e.g., *is*, *has*), there is no syntactic agreement in the Vietnamese sentences. The same is true for the Chinese examples in (4b) to (6b). For this reason, Wang (2014) calls Chinese-style imposters "pseudo-imposters". For expository convenience we will keep using the umbrella term "imposter" when there is no risk of ambiguity (as Wang himself does).

Quite often, the referents of imposters are contextually determined. In (4a) and (5a) the kin terms meaning "mom" are glossed as 1SG because they are uttered by mothers to their children, but the same sentences can be uttered by children to their mothers, and in that case the same kin terms will be glossed as 2SG. There may be crosslinguistic variation as to how much imposter interpretation depends on context, but it is clear that imposters can have context-dependent person indexing (or "floating reference", see Alves 2007 for more on Chinese influence on Vietnamese pronouns), unlike textbook pronouns, whose person indices are lexically fixed (e.g., *I* can never mean "you").

### *1.2 A new type of pronominal item*

Textbook pronouns and imposters, however, are not the only types of "pronouns" out there. Observe the sentences in (7).

(7) a. ***Mị*** *thề* *là* ***mị*** *không* *soi* *cái* *bụng.* [Viet.]
Mị.1SG swear COP Mị.1SG NEG zoom.into CLF belly
"Mị (= I) swear Mị (= I) didn't zoom into (your) belly."

b. ***Mị*** *không* *hiểu.* *Các* *chị* *hiểu* *hông?*
Mi.1SG NEG understand PL sister understand NEG
"Mị (= I) don't understand. Do sisters (= you) understand?"

In (7) the Vietnamese word *Mị* is used in place of a first-person pronoun. *Mị* is originally the name of a character in an old literary work (*Vợ chồng A Phủ* 'Couple A Phủ'), who suffered a lot of injustice in the old days but stood up for herself and fought for her own happiness. Perhaps inspired by her story, contemporary netizens (mainly females, but occasionally also males) sometimes use her name as a term of self-address with a joking tone. In (7a), for instance, the netizen says "*Mị* swear *Mị* don't feel upset" instead of "I swear I don't feel upset" when commenting on her growing belly size; and similarly, in (7b), the netizen says "*Mị* don't understand" instead of "I don't understand" to deliberately sound naïve.

There are a number of reasons why items like *Mị* differ from imposters. First, although *Mị* is originally a personal name, its pronominal usage in (7) is clearly different from that of the personal name *Linh* in (6a). Specifically, *Linh* can refer to either the speaker or the addressee whereas *Mị* can only refer to the speaker. Second, when *Linh* is used pronominally, there is actually someone named Linh in the discourse, whereas *Mị* could refer to any speaker, similarly to "I" in English. *Mị* thus behaves more like a textbook pronoun than an imposter, except that it carries idiosyncratic extragrammatical and pragmatic effects and is mainly restricted to online usage.

[3] As Paul Sidwell pointed out, it is important to note that much of the standard Vietnamese pronoun system was historically replaced by imposters.

[4] Between Vietnamese and Chinese, imposters are used even more freely in Vietnamese.

This phenomenon is not exclusive to Vietnamese but also observed in Chinese. See (8) for instance.

(8) a. *Jiù bù gěi* ***āijiā*** *xīnlǐ jiànshè* [Mandarin]
just not give mourner.1SG psychological construction
*shíjiān. Zhè jiù qīn-shàng le. Méi-rén gàosu* ***āijiā***
time now just kiss-up CRS no-person tell mourner.1SG
*háiyǒu qīn dì'èr-huí a.*
still kiss second-round EXCL
"They just don't give mourner (= me) any time to be psychologically ready. They just begin to kiss right away. No one told mourner (= me) that they were going to kiss again." (Sina Weibo[5])

b. *Guǒrán shùxué de shìjiè shì méi-yǒu sècǎi de.*
as.expected math POSS world is not-have color NMLZ
***Āijiā*** *fá le.*
mourner.1SG tired CRS
"The world of math is colorless just as expected. Mourner (= I) is tired."

Here the item of interest is *āijiā*, an ancient term of self-address that literally means "mourner" and was originally used by empress dowagers (i.e., emperors' mothers). Contemporary netizens (mostly young females, occasionally also males) often use it in a jocularly arrogant tone. Thus, the two speakers in (8) respectively complain about an unexpected kissing scene on TV and the difficulty level of math, both sounding assertive and much more fun than if the default 1SG pronoun *wǒ* is used.[6] Note that there is an ongoing debate as to whether empress dowagers in Chinese history had really called themselves *āijiā* or this was just a coinage of ancient playwrights, who then passed it on to modern scriptwriters (see Chen 2009). In spite of this, however, there is no doubt that the online term *āijiā* is borrowed from historical contexts and that it synchronically behaves like a 1SG pronoun with special pragmatic effects, similarly to Vietnamese *Mị*.

In sum, pronominal items like Vietnamese *Mị* and Chinese *āijiā*, which represent a fashionable linguistic phenomenon in the Internet era, constitute a unique category (see more examples in §2). Unlike imposters, they do not have flexible, context-dependent person indexing or common R-expression usage. Unlike textbook pronouns, they are not exponents of formal features but carry idiosyncratic extragrammatical effects. We dub them "noncanonical pronominal items" as a working term. Furthermore, noncanonical pronominal items also differ from imposters in terms of their history, typology, and lexical materials. We summarize these distinctions in Table 1.

***Table 1:*** *Differences between imposters and noncanonical pronominal items*

| | **Imposters** | **Noncanonical pronominal items** |
|---|---|---|
| Usage | wide in real life | limited to certain registers |
| History | long | emerging (mainly online) |
| Typology | prevalent in many languages | available in far fewer languages |
| Reference | flexible (contextual) | fixed (lexical) |
| R-expression usage | yes | no |
| Lexical material | nouns in contemporary use | miscellaneous |

[5] Sina Weibo is the Chinese counterpart of Twitter.

[6] Also notice the speaker's choice of verb in (7b) in the vicinity of *āijiā*. Here, for "tired", the ancient-sounding *fá* is used instead of the synchronically more common *lèi*. This sort of stylistic or register-based agreement is commonly observed in Chinese (see Feng 2010 et seq.).

As Table 1 shows, imposters have been in use in both Vietnamese and Chinese for a long time, whereas noncanonical pronominal items have only recently emerged in the Internet era. Imposters are attested in many languages (see Collins and Postal 2012 and Collins 2014), whereas noncanonical pronominal items to our knowledge are less prevalent. Finally, imposters always have contemporary nominal counterparts, which allows them to be used as ordinary R-expressions, whereas noncanonical pronominal items have miscellaneous lexical sources (e.g., *Mị* recycles a fictional character name, and *āijiā* recycles an ancient term of self-address). These distinctions set noncanonical pronominal items apart from imposters.

Despite their particular characteristics, however, noncanonical pronominal items have not been well documented. The aim of our article is thus to present a preliminary investigation of these items in contemporary usage and situate them in modern syntactic theory. To do so, we will first present noncanonical pronominal items in Vietnamese and Chinese in more detail (§2), then refine the crosslinguistic pronominal item taxonomy (§3), and eventually incorporate noncanonical pronominal items in the syntactic theory of pronouns (§4). Finally, we conclude with a few points for future research (§5).

## 2. Subtypes of noncanonical pronominal items

Noncanonical pronominal items in Vietnamese and Chinese fall in three subtypes based on their lexical sources: revived ancient terms (§2.1), dialectal terms (§2.2), and creative online coinages (§2.3).

### *2.1 Subtype I: Revived ancient terms*

The first subtype of noncanonical pronominal item we have identified in both Vietnamese and Chinese is that of revived ancient terms. Such terms can be either literary, like Vietnamese *Mị* 'female character name.1SG', or (quasi) royal, like Chinese *āijiā* 'mourner.1SG'. Table 2 contains more examples belonging to this subtype. We provide literal glosses in single quotes and original/historical usage restrictions in parentheses.

***Table 2:*** *Subtype-I noncanonical pronominal items in Vietnamese and Chinese*

| | |
|---|---|
| **Vietnamese** | *Mị* 'female character name.1SG' |
| | *trẫm* 'emperor.1SG' (by emperors) |
| | *ái-phi* 'beloved-concubine.2SG' (by emperors or princes to their concubines) |
| **Chinese** | *āijiā* 'mourner.1SG' (by empress dowagers) |
| | *zhèn* '1SG' (by emperors) |
| | *guǎ-rén* 'lacking-person.1SG' (by pre-Qin state rulers) |
| | *běn-gōng* 'this-palace.1SG' (by emperors' sons and wives/concubines to inferiors) |
| | *chénqiè* 'slave.1SG' (by emperors' wives/concubines to superiors) |
| | *qīng* '2SG' (by emperors to royal officials or between husbands and wives) |

Vietnamese *trẫm* and *ái-phi* have respectively been borrowed from Chinese *zhèn* and *ài-fēi*,[7] and have become increasing popular via hugely successful TV series such as *My Fair Princess*. *Zhèn* had originally been a default 1SG pronoun in Old Chinese (9a) but got reserved for emperors in Qin dynasty

[7] However, these cognates have clearly developed different uses in the two languages. First, Chinese *zhèn* has stayed a purely pronominal item even after being reserved for emperors, whereas Vietnamese *trẫm* seems to be in the process of further lexicalization (hence our different glosses for them). An informal survey reveals that Vietnamese speakers tend to think *trẫm* means "emperor", even though the term has no common R-expression usage (otherwise it would be an imposter like *thầy* 'teacher'). Second, while Vietnamese *ái-phi* is a noncanonical pronominal item, Chinese *ài-fēi* is an imposter (hence its absence from Table 2), for it has common R-expression usage, as in *shénmì wángye de* ***ài-fēi*** 'the beloved concubine of the mysterious prince' (novel title).

(221 B.C.E.) (9b). In the pre-Qin era, vassal state rulers humbly referred to themselves as *guǎ-rén* (9c), which literally means "a person who lacks virtue".[8]

(9) a. ***Zhèn*** *huángkǎo* *yuē* *bóyōng.* [Old Chinese][9]
1SG ancestor is.called Bóyōng
"My ancestor's name is Bóyōng." (*The Lament*, 3rd century B.C.E.)

b. *Tiān-zǐ* *zì-chēng* *yuē* ***zhèn***.
heaven-son self-refer is.called 1SG
"The Heaven's Son calls himself zhèn." (*Records of the Grand Historian*, 1st century B.C.E.)

c. ***Guǎ-rén*** *suī* *sǐ,* *yì* *wú* *huǐ* *yān.*
lacking-person.1SG even.if die also not.have regret in.it
"Even if lacking-person (= I) dies, I will have no regret." (*Commentary of Zuo*, late 4th century B.C.E.)

As we mentioned earlier, although there are debates over whether some of the ancient terms that are being revived online had really been used in history,[10] the historical origin/usage of a revived term is orthogonal to its synchronic categorial identification. It thus suffices to identify a term as a Subtype-I noncanonical pronominal item based on just two criteria: (*i*) the term has been borrowed from historical contexts (either real-life or fictional), and (*ii*) it synchronically qualifies as a noncanonical pronominal item. It is based on these criteria that we have identified the items in Table 2. See (10) and (11) for some real-life examples. Unless otherwise specified, all our Vietnamese examples are taken from Facebook or Twitter, and our Chinese examples, from Sina Weibo.

(10) a. *Hôm-nay* *có* *ai* *muốn* *rủ* ***trẫm*** *đi* *uống* [Viet.]
today have who want invite emperor.1SG go drink
*cà-phê* *không?*
coffee NEG
"Does anyone want to invite emperor (= me) out for a coffee today?"

b. ***Trẫm*** *tha* *tội* *haha.*
emperor.1SG forgive wrongdoing haha
"Emperor (= I) forgives (you) haha."

[8] It was common practice for ancient Chinese rulers to use humble terms of self-address, so in this regard *zhèn* is an exception, since it sounds authoritative and ruler-like even in archaic Chinese contexts (see, e.g., *Oracle Bone Script Dictionary* by Zhongshu Xu). The first emperor of Qin was responsible for the official royalization of the term according to official historical records (e.g., *Records of the Grand Historian*) and authoritative dictionaries (e.g., *Kangxi Dictionary*, *Xinhua Dictionary*, *Big Dictionary of Chinese Characters*).

[9] We present historical Chinese examples with Mandarin pronunciation for expository convenience.

[10] A quick search in the Chinese Text Project database (the largest online database of premodern Chinese texts) returns no results for *āijiā* (while *zhèn* and *guǎ-rén* both occur many times), and some modern dictionaries (e.g., *Revised Mandarin Chinese Dictionary*) explicitly mark *āijiā* as a term from traditional Chinese opera.

(11) a. *Jīntiān fāxiàn-le yì-gēn bái húzi,* ***zhèn*** *hěn yōushāng.* [Mandarin]
today discover-PRF one-CLF white beard zhen.1SG very sad
"Noticed a gray one in the beard today. Zhèn (= I) is very sad."

b. *Zìcóng bèi diào-dào xīn bùmén...* ***zhèn*** *jiù méi zhǔndiǎn*
since PASS transfer-to new department... zhèn.1SG still not.have on.time
*xiàbān-guò.*
knock.off-EXP
"Ever since being put in the new department, zhèn (= I) has never been able to knock off on time."

c. *Suīrán* ***guǎ-rén*** *shì dānshēn yìzú, dànshì gū*
although lacking-person.1SG is single community but lone.1SG
*juéde, nǚshēng zuì xūyào de shì péibàn.*
think girl most need NMLZ is company
"Although lacking-person (= I) is single by choice, lone (= I) thinks what girls desire is company."

d. ***Guǎ-rén*** *chí-zǎo huì sǐ-zài shèyǒu yǒngyuǎn chǎo-bù-xǐng*
lacking-person.1SG late-early will die-at roommate forever wake-not-awake
*zìjǐ de nàozhōng xià.*
self REL alarm.clock underneath
"Lacking-person (= I) will sooner or later die from my roommate's alarm clock, which can never wake herself up."

In (10), the term *trẫm* in both sentences is used to convey a jokingly arrogant tone. Similarly, in (11) *zhèn* and *guǎ-rén* sound funnily bossy and a lot less sad/mad than if the default 1SG *wǒ* were used. Note that although Vietnamese *trẫm* and Chinese *zhèn* and *guǎ-rén* were all once state rulers' terms of self-address, they have a key difference: while Vietnamese *trẫm* is only used by male speakers, the two Chinese terms are used by both male and female speakers more or less equally frequently.[11] For instance, the netizens in (11a, c) and (11b, d) are respectively male and female Weibo users, and they all sound jokingly pretentious.[12]

There are also predominantly feminine terms in this subtype. See (12) for a Vietnamese example and (13) for two Chinese examples.

(12) a. ***Ái-phi*** *hôm-nay đẹp quá!* [Viet.]
beloved-concubine today pretty INTS
"Beloved-concubine (= you) look gorgeous today!"

b. *Chào* ***ái-phi****,* *nhớ* ***ái-phi*** *quá.*
Greet beloved-concubine miss beloved-concubine INTS
"Hello beloved-concubine (= you), (I) miss beloved-concubine (= you) a lot."

[11] In particular, *zhèn* has evidently been gender-neutral throughout history, which may have to do with its original status in Old Chinese as a default 1SG pronoun. Thus, the Tang-dynasty empress Wu Zetian also referred to herself as *zhèn*, as is recorded in official historical documents like *New History of Tang* (compiled in the 11th century).

[12] Interestingly, in (11c) the speaker mixes *guǎ-rén* and *gū*. The latter, literally "lone", is an alternative to *guǎ-rén* and had also been frequently used by rulers in pre-Qin China. This suggests that the revival of ancient terms of address is a quite general trend online.

(13) a. *Jiù-mìng! Kuài gěi **běn-gōng** lái yì-píng* [Mandarin]
save-life quickly give this-palace.1SG bring one-bottle
*sù-xiào-jiù-xīn-wán!*
fast-effect-save-heart-ball
"Help! Quickly bring this-palace (= me) a bottle of instant cardio-reliever pills!"

b. *Dǎ-qiú dǎ-de **běn-gōng** yí-gè gēbo cū yí-gè gēbo xì.*
hit-ball hit-RES this-palace.1SG one-CLF arm thick one-CLF arm thin
*Zěnme pò?*
how break
"Since this-palace (= I) played too much badminton, one of my arms has become much thicker than the other. How can I get rid of this?"

c. *Wǒ yào bǎ táobǎo xiè-le, kànjiàn xīn yuèjì jiù*
1SG will DISP Taobao uninstall-PRF see new Chinese.rose just
*xiǎng mǎi... **chénqiè** rěn-bú-zhù wa!*
want buy chénqiè.1SG endure-not-TEL EXCL
"I'll uninstall Taobao, as I want to buy every new Chinese rose I see… chénqiè (= I) can't help it!"

d. *Xiū-wán liǎng-tiān jià, yòu yào huíqù bān-zhuān le.*
rest-finish two-day break again must return carry-brick CRS
***Chénqiè** bù xiǎng shàngbān le.*
chénqiè.1SG not want work CRS
"After a two-day break I must return to carry bricks (an idiom for 'work') again. Chénqiè (= I) don't want to work anymore."

In (12), the Vietnamese term *ái-phi* is used in a funnily flirtatious way to refer to a female addressee. Like *trẫm, ái-phi* synchronically retains its original gender (FEM) and is used only by male speakers to female addressees. This shows virtually no deviation from the term's historical usage. The situation in Chinese, by contrast, is much more complex. Specifically, as (13) illustrates, there are at least two different feminine terms in Subtype I, *běn-gōng* (13a–b) and *chénqiè* (13c–d). The pragmatic effect of *běn-gōng* is similar to that of *āijiā*, though it sounds slightly less bossy due to the lower ranking of emperors' wives/concubines than their mothers. *Chénqiè*, on the other hand, sounds humbler and even a bit miserable due to its historical status as a term of self-address used by low-status females to their (royal) superiors.[13] Netizens are well aware of the difference between *běn-gōng* and *chénqiè*, which is reflected in the different contexts of usage in (13a–b) and (13c–d). While *běn-gōng* is used to jokingly give orders to imaginary servants (13a) or to express "worries" about one's imperfect appearance (13b) (as royal concubines typically did), *chénqiè* is used for more miserable scenarios, such as overspending (13c) or overworking (13d).

Moreover, the historical usages of *běn-gōng* and *chénqiè* were much broader than their revived usages. Historically *běn-gōng* could be used by anyone possessing a (royal) palace, including emperors' wives, high-ranking concubines, and crown princes. But its modern revival is exclusively based on the wife/concubine sense, probably due to the omnipresence of this usage in TV series. Similarly, *chénqiè* could be used in history by any low-status female when they spoke to royal superiors, including actresses, prostitutes, plebeians, emperors' wives/concubines, and even empress dowagers when they needed to sound humble (Xia 2018). But its online usage is only based on the wife/concubine sense too, again due to its omnipresence in TV series. In addition, while these terms are predominately used by females, they are occasionally also used by males (mostly gay). For instance, a flamboyant gay character Yu Hao in a TV series *Stand by Me* constantly refers to himself as *āijiā*, and a gay vlogger and cosmetics expert Benny Dong on Bilibili (the Chinese YouTube) regularly calls himself *běn-gōng* in his video

[13] The ultimate origin of *chénqiè* was an Old Chinese compound meaning "slaves" (lit. male slave [*chén*] and female slave [*qiè*]). However, this original sense had long become obsolete, and *chénqiè* shifted to its feminine usage in as early as Eastern Han dynasty (25–220 C.E.). See Xia (2018) for a detailed discussion.

titles. As for *chénqiè*, the catchphrase *Chénqiè zuò-bú-dào a!* 'I really can't do it!'[14] is trendy among netizens of all genders and sexual orientations.

A word of caution is in order on the category of *běn-gōng*. Although its literal meaning "this-palace" makes it resemble English *this reporter*, *the present author*, etc., which are imposters à la Collins and Postal, we must note that this Chinese term has no R-expression usage or flexible reference. Thus, it cannot be used in third-person cases like (14a), unlike English *this*-terms, as in (14b).

(14) a. ***Běn-gōng** bǐ bié-gōng měi.*
this-palace.3SG compared.to other-palace beautiful
"Intended: This palace (= she) is more beautiful than other palaces (= other concubines)."

b. **This reporter$_{3SG}$** is nicer than that one.

In other words, Chinese *běn-gōng* is a lexically fixed, idiomatic term of self-address, which makes it qualify as a noncanonical pronominal item in our criteria.[15]

### *2.2 Subtype II: Dialectal terms*

The second subtype of noncanonical pronominal item we have identified in Vietnamese and Chinese involves dialectal terms that have made their way into the common language via mass media (e.g., TV programs) or the Internet. See Table 3 for some examples.

***Table 3:** Subtype-II noncanonical pronominal items in Vietnamese and Chinese*

| | |
|---|---|
| **Vietnamese** | *hắn* '3SG' (from Central and Southern dialects)[16] |
| | *y* '3SG' (from Northern dialects) |
| **Chinese** | *ǒu* '1SG' (from Min/Yue Chinese) |
| | *é* '1SG' (from Shaanxi Mandarin Chinese) |
| | *ǎn* '1SG' (from Northern/Central Mandarin Chinese) |
| | *nóng* '2SG' (from Shanghai Wu Chinese) |
| | *yā* '3SG' (from Beijing Mandarin Chinese) |

There are several distinctions between dialectal terms and revived ancient terms. First, while revived ancient terms are restricted to the first and the second person, dialectal terms also involve third-person items. For Vietnamese in particular, these terms are strictly 3SG (see Alves 2017 for more detail on the etymology of these). This is not surprising because dialectal terms are simply default pronouns in their original dialects. Second, unlike revived ancient terms, dialectal terms may not have gender restrictions at all, and in some cases not even preferences. Third, unlike many revived ancient terms (especially the royal ones), dialectal terms generally do not bear arrogant or bossy tones, so their pragmatic effects are of a different sort. Consider (15) for Vietnamese.

[14] This is a quote from the highly popular TV series *Empresses in the Palace* and has gone viral via memes.

[15] In this sense *běn-gōng* patterns more like *yours truly* and *muggins here* in English, which have no R-expression usage or flexible reference either, even though they are classified as imposters in Collins & Postal (2012).

[16] Although *hắn* has been documented in various parts of Vietnam, including the North (Cao 2014), the term was originally from the Southern dialect (see, e.g., Hoang 1989). It should also be noted that while *hắn* is mostly used as a neutral 3SG in the South, it is often used with pragmatic effects in the North (addressed later in this section).

(15) a. *Đi rồi không biết bao-giờ* ***hắn*** *mới về.* [Vietnamese, Southern speaker]
go PRF NEG know when 3SG PRT return
"(He)'s gone. (I) don't know when s/he would return."

b. *Không biết thì hỏi* ***hắn*** *thử.* [Southern speaker]
NEG know then ask 3SG try
"If (you) don't know then try asking him/her."

c. *Chính* ***hắn*** *cục vàng của tui.* [Northern speaker]
precisely 3SG CLF gold POSS 1SG
"It's precisely him, my piece of gold."

d. *Lại nhớ* ***hắn*** *à?* [Northern speaker]
again miss 3SG Q
"Are (you) missing him again?"

e. *Mèo nhà tao đã bắt* ***y*** *sáng nay rồi*. [standard variety, Internet language]
cat home 1SG PST catch 3SG morning DEM PRF
"My cat caught it this morning."

f. *Tôi nói lời yêu* ***y****, nhưng sao* ***y*** *không hiểu?* [standard variety, Internet language]
1SG say word love 3SG but why 3SG NEG understand
"I said loving words to him, but why hasn't he understood?"

First, the term *hắn* is specific to Central and Southern dialects of Vietnamese, is a regional variant of the standard 3SG *nó*, and has entered the standard variety due to dialect contact. Note that although *hắn* is a neutral term in the original dialect, as in (15a–b), it is often used to sound cute/funny by speakers of other varieties. In (15c), for example, the speaker of Northern Vietnamese uses *hắn* as an endearing term to refer to her baby, who she considers "a piece of gold" in her possession.[17] Similarly, another Northern Vietnamese speaker in (15d) uses *hắn* to refer to their interlocutor's boyfriend, who is presumably being missed by the interlocutor. The use of *hắn* in both contexts sounds more fun and cuter than the standard variant *nó*.

The next Vietnamese item in this subtype is *y*, which is originally and mainly used in Northern dialects to refer to a male criminal. The term is therefore formal, but due to crossdialectal contact it has now become more widely used online as a jokingly serious pronominal form. In (15e), for example, *y* is used to refer to a mouse (who is in this sense cast as a criminal), which makes the sentence much funnier. Similarly, the speaker in (15f) uses *y* to complain about her crush, who has not returned her affection. Like *hắn,* the use of *y* in these contexts brings about some dramatic comic effects. Unlike *hắn,* however, there is a strong preference for a masculine interpretation particularly when the referent of *y* is human. When the referent is nonhuman, *y* can in principle be used neutrally (e.g., in (15e), we do not know whether the mouse is male or female).

The Chinese inventory for this subtype is again more diverse. Due to space limitations, we restrict our detailed description to only three of the Chinese terms from Table 3. All examples in (16) are from Common Mandarin produced by Weibo users.

[17] *Cục vàng* 'a piece of gold' is an idiomatic expression in Vietnamese which is most often used by parents to refer to their precious children.

(16) a. ***Ǒu*** *zhè-ge yǎn-zhū-zi hǎo-kàn.* [Mandarin]
1SG this-CLF eye-ball-DIM good-look
"My eyes are good-looking."

b. *Wèishénme* ***ǒu*** *de xīn-dòng nánshēng hái bù chūxiàn?*
why 1SG POSS heart-move guy still not appear
"Why hasn't the guy I will fall in love with shown up yet?"

c. ***É*** *dìyī-cì pèngdào chǎo-wán jià jǔbào jiā gào-lǎoshī de.*
1SG first-time encounter argue-finish quarrel report plus tell-teacher NMLZ
"This is the first time I have ever encountered someone who reports the other person to teachers after quarreling with them."

d. ***É*** *di guīmì jìngrán bǎ* ***é*** *zuì-ài de*
1SG POSS best.female.friend go.so.far.as.to DISP 1SG most-love REL
*xiāngshuǐ yòng-zuò kōngqì-qīngxīn-jì!*
perfume use-as air-freshen-agent
"My bestie outrageously used my favorite perfume as air freshener!"

e. *Nǐ hé zhǔrén shuō yí-jù ràng tā shuān shéng,* ***yā*** *zhāng-kǒu*
2SG with owner say one-CLF let 3SG tie rope 3SG.OFF open-mouth
*jiù xiàoxīxī hé nǐ shuō: "wǒ jiā gǒu bù yǎo rén."*
just giggle with 2SG say 1SG house dog not bite person
"You ask the dog owner to tie their dog up, and they$_{3SG}$ (= that asshole) just giggle and tell you: 'My dog does not bite.'"

f. *Nǐ yuè gēn* ***yā*** *shuō hǎo-tīng-de* ***yā*** *yuè lái-jìn.*
2SG the.more with 3SG.OFF say good-listen-NMLZ 3SG.OFF the.more come-strength
*Yíshànglai bǎ* ***yā*** *mà-xiaqu... hái néng yǒu diǎn yòng.*
right.at.the.beginning DISP 3SG.OFF scold-down still can have some use
"The more kind words you say to them$_{3SG}$ (= that asshole), the more shameless they$_{3SG}$ (= that asshole) become. It would be more effective if you simply swear back and tell them$_{3SG}$ (= that asshole) to get out of the car right away."

The three terms in (16) have respectively been borrowed from Min/Yue Chinese (*ǒu*), Shaanxi Mandarin Chinese (*é*), and Beijing Mandarin Chinese (*yā*). The first item, *ǒu*, is a regional variant of Mandarin *wǒ* and (re)entered Common Mandarin due to netizens' mocking of the dialectal pronunciation. According to Chen (2009:215), it was the most popular mutant personal pronoun online in the noughties. Perhaps due to its initial role as a mocking term, *ǒu* sounds funny and cute and is often used by netizens who want to appear jolly and adorable. The current usage of *ǒu* is no longer for mocking purposes. Thus, the netizen in (16a) happily posts about her satisfaction with the look of her eyes without the intention to mock anyone, and the netizen in (16b) laments her single status with a puppy-face-like tone. In both sentences, the use of *ǒu* instead of the default 1SG *wǒ* makes the speakers sound more likable and less boastful or whiny.

The second item, *é* (sometimes rendered as *è* or *ngè*) became widely known in the noughties via the popular TV series *My Own Swordsman*, in which the leading actress spoke Shaanxi Mandarin throughout the eighty episodes. Due to the comedic nature of that show and the fussy personality of its main character, the term has subsequently gained a jokingly fussy tone in Internet language. Thus, the netizen in (16c) is making a fuss about the base behavior of a tattletale student, and that in (16d), about her best friend's inadequate use of her perfume. No such fussy tone would be present if the default 1SG *wǒ* were used instead. In addition, in (16d) the possessive marker *di* is also borrowed from Shaanxi Mandarin, whose Standard Mandarin counterpart is *de*. This again reflects the stylistic agreement mentioned in footnote 6.

A special note is in order concerning the Beijing Mandarin term *yā*, which is originally a highly vulgar expression meaning "child of a girl with no recognized marital status". It is etymologically short for *yātou-yǎng-de* 'low.status.girl-raise-NMLZ', but nowadays this literal meaning is obsolete, and *yā* is

mainly used as an offensive suffix attached to pronouns and demonstratives (e.g., *nǐ-yā* '2SG-OFF', *nèi-yā* 'that.person-OFF').[18] That said, it has developed a stand-alone pronominal usage as well. Crucially, in this usage, it can only be interpreted as 3SG (gender neutral). Thus, the netizens in (16e–f) respectively complain about the bad behavior of a dog owner and that of a taxi customer. Since *yā* has no synchronic R-expression usage related to its pronominal usage[19] and has a lexically fixed person index, we treat it as a noncanonical pronominal item instead of an imposter.

### *2.3 Subtype III: Creative online coinages*

The third subtype of noncanonical pronominal item we have identified in Vietnamese and Chinese involves items that do not fall in the previous two subtypes. These are mainly creative coinages on the Internet and so in a sense "native" to Internet language. Since such online coinages are not based on any particular type of source, their lexical materials are miscellaneous or even totally novel. See Table 4 for some examples.

Note that all three Vietnamese examples in Table 4 are second-person terms, which carry different pragmatic effects due to the lexical materials they recycle. See (17) for some real-life examples.

***Table 4:*** *Subtype-III noncanonical pronominal items in Vietnamese and Chinese*

| | |
|---|---|
| **Vietnamese** | *cưng* 'dear.2SG' |
| | *con-quỷ* 'devil.2SG' |
| | *người-đẹp* 'beautiful.person.2SG' |
| **Chinese** | *qīn* 'dear.2SG' |
| | *běn-lū/běn-lú/běn-lǔ* 'this-loser.1SG' |
| | *lúnjiā* 'others.1SG' |

(17) a. ***Cưng*** *muốn gì từ anh nào?* [Viet.]
dear.2SG want what from 1SG.MASC AFFECT
"What do you (= dear) want from me?"

b. ***Con-quỷ*** *đang làm gì đó?*
devil.2SG PROG do what DM
"What are you (= devil) doing?"

c. *Cảm-ơn* ***người-đẹp*** *đã mở hàng.*
thank beautiful.person.2SG PST open shop
"Thank you (= beautiful person) for being the first customer today!"

All three boldfaced terms in (17) are creative coinages by market sellers, which are now widely used thanks to online marketing. Similar to noncanonical pronominal items in the other two subtypes, those in Subtype III encode special pragmatic effects too. For example, the uses of *cưng* and *con-quỷ* as 2SG terms in (17a–b) sound deliberately cute and friendly (and possibly a little flirtatious), while the use of *người-đẹp* in (17c) is flattery/fashionable. These terms are very creative and have no fixed lexical sources. Similarly, see (18) for some real-life examples of Chinese terms in this subtype.

[18] A *de* can be optionally added to these terms (e.g., *nǐ-yā-de*), which is a residue of the nominalizer in the full form.

[19] The character for *yā* (丫) does record other meanings too, such as "branch, twig" or more generally any Y-shaped object like the front part of a foot, but those are irrelevant to the pronominal *yā*. Thus, such polysemy is qualitatively different from that in cases like "teacher" as an R-expression and "teacher" as a term of address.

(18) a. ***Qīn*** *kěyǐ shōucáng wǒ-men de diànpù hé liànjiē...* [Mandarin]
dear.2SG can save 1SG-PL POSS store and link
*xūyào de shíhòu zài liánxì ne.*
need REL time again contact EMP
"Dear (= you) can save our store and link and contact us again when you have need."

b. *Suīrán* ***běn-lú*** *méi qián méi míng, dàn wǒ ānquán*
though this-loser not.have money not.have fame but 1SG safety
*yìshì yīliú.*
awareness first-class
"Although this loser (= I) has no money or fame, I have first-class safety awareness."

c. ***Lúnjiā*** *zhēndeshì huīcháng huīcháng xǐhuān máo búyì o.*
others.1SG really very very like Mao Buyi EXCL
"Others (= I) really like Mao Buyi very very much!"

*Qīn* originated on the shopping website Taobao as a friendly term of address between sellers and customers. Thus, in (18a) the Taobao seller says "dear can…" instead of "you can…" to encourage the customer to save their online store. *Běn-lú* and its tonal variants are all coined by combining the deictic *běn* 'this' and the first syllable of English *loser*. It has the pragmatic effect of self-mocking.[20] Thus, in (18b) the netizen self-mockingly takes pride in his safety awareness despite his poor status. *Lúnjiā* is a deliberately distorted variant of *rénjiā* 'others' and sounds cute and jocular when used as a 1SG term. Thus, in (18c) the netizen expresses her obsession with the pop singer Mao Buyi. Also note the form *huīcháng* in (18c), which is a distorted variant of the degree adverb *feīcháng* 'very' and adds to the cuteness of the utterance. This is another instance of the aforementioned stylistic agreement (see footnote 6).

Since this is the first systematic documentation of such creative pronominal coinages to our knowledge, we want to give a bit more detail on the above terms to justify our identification of them as noncanonical pronominal items. First, we are aware that *qīn* has R-expression usages, partly due to the versatility of its lexical root, which can mean "parent, kin (n.)", "intimate, dear (adj.)", "kiss (v.)", etc. Moreover, one of its R-expression usages is closely related to its pronominal usage. Thus, one can friendly refer to someone (a third person) as *qīn*, as in (19).

(19) *Wēibó hái yǒu* ***qīn*** *zài ma?* [Mandarin]
Weibo still have dear.N be.at Q
"Are there still anyone$_{\text{friendly}}$ on Weibo?"

However, we do not treat *qīn* as an imposter because its pronominal usage has evidently developed from its term-of-address usage, which is also its predominant usage on Taobao (see, e.g., Deng 2012 and Liu 2012). In fact, previous studies rarely mention the usage exemplified in (19), which suggests that it might be a more recent development from either the pronominal or the term-of-address usage. In any event, the 2SG *qīn* is not a pronominally used R-expression in nature and thus does not fit the canonical situation of imposters.

Second, just like the Subtype-I term *běn-gōng* 'this palace.1SG', *běn-lū* also contains a deictic *běn*, but we treat it as a noncanonical pronominal item instead of an imposter because it too can only refer to the speaker (but not a third person) and has no R-expression usage, as exemplified for *běn-gōng* in (14).

Third, the phonological distortion that has created *lúnjiā* (*rén→lún*) has brought along some interesting change to its syntactic status. On the one hand, while both *rénjiā* and *lúnjiā* can be used as 1SG terms, only *rénjiā* has a separate 3SG usage (i.e., "others"), hence the ambiguity of (20a). By contrast, *lúnjiā* can only refer to the speaker, hence the ungrammaticality of the 3PL reading in (20b).

[20] This term, especially its *běn-lū* variant, is mainly used by males, since the Chinese character usually adopted to represent *lū* (撸) also means "(male) masturbate", which further adds to the self-mocking effect of the term.

(20) a. *Bié zhuāng le!* ***Rénjiā*** *yòu bú shì shǎzi.* [Mandarin]
don't pretend CRS others.3PL/1SG EMP not COP idiot
"Stop pretending! Others (= they/I) are not idiots."
b. *Bié zhuāng le!* ***Lúnjiā*** *yòu bú shì shǎzi.*
don't pretend CRS others.*3PL/1SG EMP not COP idiot
"Stop pretending! Others (= *they/I) are not idiots."

On the other hand, while *rénjiā* has a shy or embarrassed tone when used as a term of self-address, *lúnjiā* furthermore sounds adorable and cartoon-like. Thus, while (20a) sounds like real blaming (with an embarrassed tone in the 1SG reading), (20b) sounds like the speaker is just teasing the addressee.

### *2.4 Interim summary*

Based on Vietnamese and Chinese data, we have identified and exemplified a new type of pronominal item that is emerging in the Internet era, which we have dubbed noncanonical. As we have shown, noncanonical pronominal items differ from both textbook default pronouns and imposters in nontrivial ways. Crucially, unlike imposters, they have lexically fixed referents and no common R-expression usage, and unlike default pronouns they have various pragmatic effects. Overall, though, they pattern almost identically to default pronouns in syntax except for their extragrammatical effects. So, we tentatively rename noncanonical pronominal items "noncanonical pronouns" and give them the following working definition:

(21) **Noncanonical pronouns** are syntactically well-behaved pronouns with extragrammatical effects.

Moreover, the extragrammatical effects in noncanonical pronouns are not associated with conventional sociolinguistic factors like the relationship or relative hierarchical status between speakers and addressees.[21] In fact, most examples we have given are not even from interpersonal communication but from online posts. Rather, the extragrammatical (e.g., register, tone) effects we have observed are more typically associated with the mind-sets and personalities of individual netizens themselves. To illustrate, adjectives we have used to describe the special effects of Vietnamese and Chinese noncanonical pronouns include the following:

(22) joking, jocularly arrogant, fun, funnily bossy, jokingly pretentious, funnily flirtatious, miserable, cute, jokingly serious, dramatic, jolly, adorable, puppy-face-like, jokingly fussy, offensive, deliberately cute, friendly, flattery, fashionable, self-mocking, cartoon-like, teasing

These descriptions are highly compatible with the Internet register. For instance, while it would sound bizarre or even off-putting if an adult keeps talking in a dramatic, cute, or cartoon-like tone in reality, this is totally fine and acceptable on the Internet. In a similar vein, while gender does play a role in regulating the use of noncanonical pronouns, at least in Chinese, it is apparently one's gender identity (or sexual orientation) rather than their biological sex that guides their choices of noncanonical terms. This is another state of affairs increasingly normal in contemporary Chinese society, especially on the Internet.

Perhaps due to the unique features of the Internet as a modality of communication and the somewhat similar technological context it has endowed netizens around the world with, we have observed striking similarities in Vietnamese and Chinese noncanonical pronouns not only in their usage but also in their lexical sources. Specifically, for both languages we have identified three major types of noncanonical pronouns based on their evolution pathways: revived ancient terms, dialectal terms, and creative online coinages. Also, for certain terms (e.g., *trẫm*, *ái-phi*), we have even observed mass-media-based crosslinguistic borrowing.

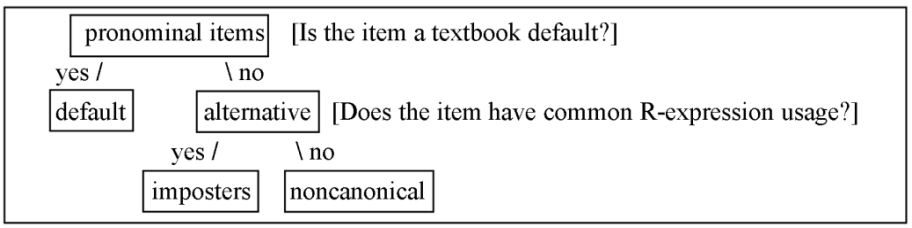
[21] So noncanonical pronouns are different from the kind of interpersonal-relationship- or social-hierarchy-based pronominal items typically seen in East and Southeast Asian languages like Japanese, Korean, and Thai.

On the other hand, Vietnamese and Chinese show two main differences in their noncanonical pronouns. First, they differ in the sizes of their respective inventories, with Chinese having more actively used terms in almost every subtype. Second, they differ in the person/gender/number propensities of certain terms or even entire subtypes. For instance, Vietnamese terms like *trẫm* (male) and *ái-phi* (male to female) show rather strict gender-based usage, and both Subtype-II and Subtype-III terms in Vietnamese show person/number restrictions (to 3SG and 2SG respectively). Whether these restrictions are categorical or due to the limited size of our data set requires further investigation, but our observation so far suggests that while Vietnamese has a more developed and less restricted imposter system than Chinese (see footnote 4), Chinese has a more developed and less restricted noncanonical pronoun system than Vietnamese. What this contrast means is an intriguing point of future research, but in the rest of this article we focus on the formal-grammatical status and syntactic representation of noncanonical pronouns.

## 3. Taxonomy

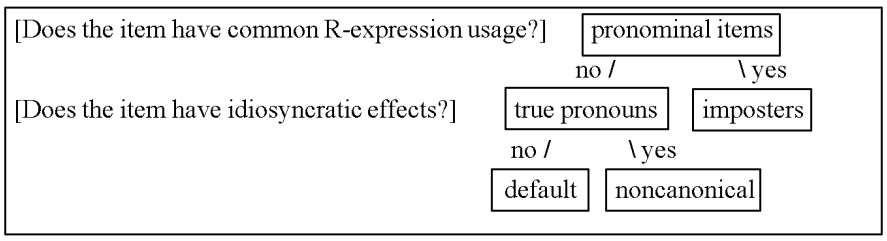
Before presenting our formal syntactic analysis, we first make a brief remark on the taxonomy of pronominal items. At first sight, an intuitive way of classifying pronominal items is to build on the conventional textbook system. Since textbook pronouns are deemed standard or default, nontextbook pronominal items could be termed alternative or nondefault. Then, under the alternative class, we further set two subclasses: imposters and noncanonical pronouns. We illustrate this taxonomy with the diagram in Figure 1 and call it the usage-based taxonomy, since properties like conventional/nonconventional and canonical/noncanonical are from the perspective of language use. Also, to give the taxonomy a bit more systematicity, we recast the two binary-branching nodes in terms of two yes-no questions.

***Figure 1:*** *A usage-based taxonomy of pronominal items*




However, from a grammatical perspective the taxonomy in Figure 1 is obviously flawed, because even though imposters and noncanonical pronouns are both nondefault pronominal items, grammatically speaking noncanonical pronouns resemble default pronouns to a much greater extent than they resemble imposters. So, in a grammatically more precise taxonomy default and noncanonical pronouns should be grouped under the same class at some level. To this end, we propose the syntactically based taxonomy in Figure 2, which is also accompanied by two yes-no questions defining the two binary-branching nodes.

***Figure 2:*** *A syntactically based taxonomy of pronominal items*




Incidentally, the syntactically based taxonomy is also more desirable from an acquisitional perspective, since the two questions in Figure 2 can both be easily answered based on the primary linguistic data a child has direct access to, whereas the first question in Figure 1 hinges on more sophisticated knowledge about the world, more exactly about language textbooks and reference grammars. According to the

principles and parameters approach to linguistic variation and diversity (Chomsky 1981), especially its neo-emergentist incarnation (Biberauer 2017), children acquire grammatical knowledge based on positive evidence in the significant acquisitional input (e.g., high-frequency recurring forms and collocations in everyday adult speech). So, a certain grammatical phenomenon is acquired early (or at least is acquirable) if it regularly and saliently exists in the primary linguistic data and therefore can be easily detected by the acquirer. Building on this theoretical background, a further prediction Figure 2 makes is that in languages where imposters are a regular part of the grammar, such as Vietnamese, imposters are acquired earlier than default pronouns.[22] We leave the verification of this claim to future research.

## 4. A formal syntactic analysis

After laying out the comparative data (§2) and the syntactically based taxonomy (§3), in this section we propose a formal syntactic analysis of noncanonical pronouns within the minimalist program (Chomsky 1995 et seq.), more exactly within the generalized root syntax theory put forth in Song (2019). The purpose of this analysis is threefold: (i) to explain why noncanonical pronouns behave the way they do in a formally explicit way, (ii) to improve current syntactic theory of pronouns and make it empirically more adequate, and (iii) to tentatively explain why noncanonical pronouns have restricted crosslinguistic distribution. Before presenting our particular analysis (§4.2), we first introduce the theoretical background of pronominal syntax (§4.1).

### *4.1 Pronouns in generative syntax*

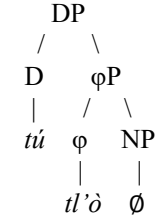
As we mentioned in §1, previous syntactic studies of pronouns have mainly focused on default pronouns, especially those that constitute exponents of phi features. Within the Chomskyan school, since the proposal of the DP hypothesis (Abney 1987), default pronouns have been associated with the category D (or its more elaborate equivalents) in one way or another. Thus, Abney (building on Postal 1966) puts pronouns at the D head position and argues that they project DPs on their own (i.e., without specifier/complement elements). Then, with the popularization of the split-functional-projection idea, which was initiated by Pollock's (1989) split-IP hypothesis and Rizzi's (1997) split-CP hypothesis in the verbal domain, many authors have proposed elaborate hierarchical structures for the nominal domain as well (e.g., Ritter 1995, Cardinaletti and Starke 1999, Neeleman and Weerman 1999, Borer 2005, Ritter and Wiltschko 2019). In relation to pronouns, quite a few authors have put forth the idea that they may correspond to different parts or zones in the nominal tree.

For instance, Ritter (1995) proposes two kinds of pronouns, which respectively occupy D and Num. Déchaine and Wiltschko (2002) propose three kinds of pronouns, which they name pro-DP, pro-φP, and pro-NP. Like Ritter, Déchaine and Wiltschko also let pronouns occupy head positions. But that is not the only solution. There are also researchers who propose that pronouns realize entire nominal tree(let)s (e.g., Weerman and Evers-Vermeul 2002, Neeleman and Szendrői 2007). As Neeleman and Szendrői suggest, this approach is more natural in late spell-out frameworks like distributed morphology (Halle and Marantz 1993 et seq.), in which syntactic computation operates on formal features, whose phonological realization is only dealt with at the syntax-phonology interface. Since such a computation-before-pronunciation view is also more generally endorsed in the minimalist program, we accept it as a background assumption without further discussion. That is, we assume that both terminal and nonterminal nodes may be spelled out by phonological units. This means that what looks like a single word or even morpheme on the surface may be a complex hierarchical structure in the underlying syntax and that generative syntax can well handle this type of phenomenon (by means of nonterminal spell-out).

Special attention needs to be paid to two recurring issues in previous generative analyses of pronouns. The first issue is the division of labor in the elaborate structure of pronouns. Take Déchaine and Wiltschko's pro-DP structure for example.

[22] However, Figure 2 does not predict that default and noncanonical pronouns are acquired simultaneously, because the latter are a novel phenomenon in Internet language and not yet part of the grammatical knowledge relevant for first language acquisition or parameter setting.

(23) Pro-DP (illustrated by Halkomelem *tú-tl'ò* 'DET-3SG'; Déchaine and Wiltschko 2002:412)

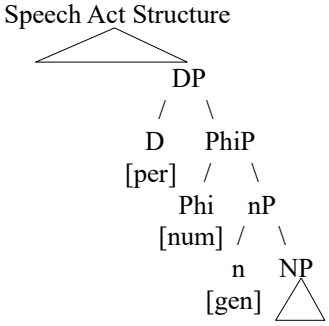



In this structure, both D (for definiteness) and φ (for person/number) are overt, but NP is empty. Déchaine and Wiltschko hypothesize that the NP position may be either overt or null, and that when it its filled with lexical material we get a normal [determiner noun] phrase, such as (24).

(24) *Tl'ó-cha-l-su qwemcíwe-t* [*thú-tl'ó q'ami*]ARG. [Halkomelem]
then-FUT-1SG-so hug-TRANS DET.FEM-3SG girl
"Then I'm going to hug that girl." (Galloway 1993:174, via Déchaine and Wiltschko 2002:412)

However, Déchaine and Wiltschko do not specify what the semantic contribution of NP is when it is null as in (23). A similar scenario occurs in Ritter and Wiltschko (2019), where an even more elaborate nominal domain is proposed, which has three zones—a lexical, a functional, and an interactional zone—each subsuming a number of categories. Take their analysis of German *du* '2SG' in (25) for example.

(25) German pronoun *du* '2SG' (Ritter and Wiltschko 2019:3)

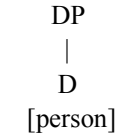



The authors do not discuss the spell-out procedure, but their theory is compatible with a late spell-out approach (Wiltschko, p.c.), so we can assume that the phonological unit *du* somehow realizes the whole tree in (25). Again, we see concrete semantic contributions of each syntactic category except NP, though it presumably has to be there since otherwise the tree is not lexically grounded.

Lexical grounding is precisely the second issue in previous syntactic studies that we would like to invite readers to pay attention to. While both (23) and (25) have a functional-above-lexical scaffolding, as is standardly assumed in current generative syntax, this is not always the case in earlier studies of pronouns. For example, the following trees from Abney (1987) and Ritter (1995) have no lexical bases (i.e., NP).

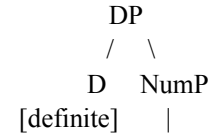
(26) a. Abney (1987:284)


b. Ritter (1995:419)




(1st/2nd-person pronouns in Hebrew)

c. Ritter (1995:421)

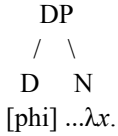



(3rd-person pronouns in Hebrew)

It is unclear how such trees can be generated in minimalism, in which functional categories by definition build on or extend lexical categories. In fact, in some minimalist theories, such as distributed morphology, the lexical grounding even provides the categorial feature for the entire tree (Roberts 2019 has a similar idea).

The reason why we want to draw readers' attention to the above two issues is because they are prevalent in the literature and deserve reflection if we want to guarantee the implementability of particular theories in the minimalist program (e.g., in the fashion of Collins and Stabler 2016). That said, we do not think the theories cited above are inherently problematic. Rather, they are just of a lower granularity level, where certain details are glossed over. As a granularity-raising move, we find an idea in Harbour (2016) helpful. Abstracting away from the technical details, Harbour's core idea is that the individual variable (a semantic primitive), which any further nominal semantic function (be it person, number, or whatever else) relies on, is introduced *before* the further formal features are introduced. He lets the individual variable be introduced at the N level and the further features be introduced at the D (or split-D) level, as in (27).

(27) Semantic division of labor in the nominal structure (adapted from Harbour 2016:77)




The overall effect of (27) is that the phi features in D together yield a particular set of individuals, which the individual variable in N ranges over.[23] Eventually, some other features in the nominal structure pin down the referent and assign it to the individual variable *x*. Acquaviva (2019) independently proposes a similar idea in an even more fine-grained theory, where the N part is further decomposed into a nominalizer *n* and a root in the sense of distributed morphology, and it is the nominalizer that introduces the individual variable that serves as the semantic grounding of the whole nominal phrase.

Assuming the above high-granularity details in the background, we can now safely take any of the aforementioned theories of default pronouns as a point of departure without worrying about shaky foundations. In fact, which specific DP structure we adopt is immaterial to our own analysis as long as it does not suffer from foundational problems and has the necessary components to derive a syntactically well-behaved default pronoun. Therefore, when the DP-internal structure is inconsequential, we simply use the very-low-granularity label $DP_{pro}$ to indicate a pronominal DP (as Aldridge 2021 does in her study on Old Chinese pronouns).[24]

Having seen how default pronouns are typically analyzed in generative syntax, we now build our analysis of noncanonical pronouns on top of that, because as we mentioned in §2.4, noncanonical pronouns are just default pronouns equipped with idiosyncratic extragrammatical effects. Since they

[23] Harbour's tree is highly abbreviatory, which cannot be taken at face value for issues like labeling (Chomsky 2013 et seq.). We assume that the tree in (27) automatically gets the label DP once the omitted details are filled back.

[24] We use DP as an umbrella label for the entire (pro)nominal structure, which abbreviates not only the lexical and functional zones but also the Wiltschkovian interactional zone if that is present.

behave like default pronouns in syntax, they must involve the same sort of syntactic structure as default pronouns, whatever that structure is. What our analysis mainly serves to explain, therefore, is how the idiosyncratic effects are formally associated with that structure.

### *4.2 Generalized root syntax*

To begin with, note that the special effects of noncanonical pronouns come from the particular terms themselves, more exactly from the idiosyncratic content associated with their lexical materials, rather than from the context. This is because they show the same effects in all examples. For instance, Vietnamese *trẫm* and Chinese *zhèn* (both 1SG) sound jokingly arrogant wherever they appear, and they have this effect precisely because they were once used by emperors. Similarly, Vietnamese *cưng* and Chinese *qīn* (both 2SG) sound deliberately friendly, either in commercial discourses or not, and this effect has to do with the terms' original lexical meaning "dear". An even more interesting case is Chinese *lúnjiā* 'others.1SG', which not only takes some lexical material (*rénjiā* 'others.1SG/3SG/3PL') but also distorts it (*rén→lún*), and the distortion in turn serves to distinguish the new noncanonical usage from the old imposter usage. The key mechanism involved in all these cases of noncanonical pronouns is thus the recycling of existing lexical material for new grammatical purposes (i.e., a kind of grammaticalization). This process is evident in all three subtypes of noncanonical pronouns in §2. It is also a manifestation of a fundamental strategy in human language and cognition, which Biberauer (2017) terms "maximize minimal means" (MMM). In Biberauer's (2017:41) words, MMM is both "a generally applicable learning bias harnessed by the acquirer during acquisition" and "a principle of structure building" that "facilitat[es] the kind of efficient computation and … the self-diversifying property that allows human language to be the powerful tool that it is" (see Biberauer 2011 et seq. for more background on this line of thought).

Song (2019) develops a "generalized root syntax" to tackle half-grammatical-half-lexical vocabulary items, of which the noncanonical pronouns studied here are a specific instance. Song's theory is an extension of the root theory in distributed morphology (hence its name), where content words like nouns and verbs are decomposed into a functional part, called the categorizer, and a purely lexical part (which does not have a syntactic category), called the root. Thus, the noun *dog* is analyzed as [*n* √DOG], and the verb *run* is analyzed as [*v* √RUN]. The idea is that all that participates in formal computation is essentially functional-categorial when the representation is fine-grained enough, while idiosyncratic information like lexical sound/meaning and encyclopedic knowledge is sealed in a syntactically inert capsule that is only opened when the syntactic representation is sent to the phonological/semantic interfaces for interpretation. This lexical decompositional practice pushes syntactic methods to the traditional morphological arena, so distributed morphology is also known as a "syntax all the way down" approach.[25]

While accepting the lexical decompositional approach of distributed morphology, Song points out that its particular view on root categorization is flawed, for it stipulates that only traditional lexical categories can serve as categorizers, but that assumption leads to theory-internal contradiction under close scrutiny (see Song 2019:102 for details). Since from a formal perspective the categorization procedure just serves to equip the otherwise inert root with a syntactically active shell, logically speaking any functional category can do the job, and the specialness of traditional lexical categories (i.e., the little *x* categorizers in distributed morphology) merely lies in their bare-predicate-making/typing semantics. That is, they introduce typed individual variables in the sense of Harbour (2016) and Acquaviva (2019):

[25] A caveat here is that distributed morphology does *not* predict that any root-categorizer merger can yield a legitimate vocabulary item (which is a common straw man in criticism of the framework). Rather, the interpretability of particular root-categorizer combinations is a matter of language-specific lexicalization (in a broad sense of the term), and root syntax merely offers a tool to structurally represent and analyze such lexically stored information. A biggest achievement of root syntax, in hindsight, is that it has pushed syntactic theory to a higher granularity level and thereby formalized a further aspect of regularity in human language (i.e., the very basic phenomenon of categorization). We deem this a significant step forward.

(28) a. N-categorizer ("nominalizer"): introduces an `entity`-type individual variable
b. V-categorizer ("verbalizer"): introduces an `eventuality`-type individual variable
c. X-categorizer: introduces an `x`-type function

Thus, root categorization is more generally root support; namely, the enrichment of a functional category with some idiosyncratic information encapsulated in a root. For instance:

(29) a. *dog* = [*n* √DOG] = an `entity`-type individual that is called /dɔg/, has four legs, can bark, etc.
b. *run* = [*v* √RUN] = an `eventuality`-type individual that is called /rʌn/, involves leg-moving, etc.

From the perspective of syntacticosemantic computation, only the underlined information in (29) is relevant, while the rest merely supports this skeletal information and makes it suitable for postsyntactic purposes (describing the world, communication, etc.). In other words, the extragrammatical effects of content words is precisely their idiosyncratic root content. Extending this mechanism to X-categorizers, what we obtain is a functional category equipped with idiosyncratic extragrammatical effects; namely, a half-grammatical-half-lexical item. Song (2019) gives miscellaneous examples from Chinese to illustrate this, such as those in (30).

(30) a. passive auxiliaries: *bèi* 'lit. cover, suffer (neutral)', *gěi* 'lit. give (colloquial, negative)'
b. classifiers: *jiàn* 'lit. item (for clothes, etc.)', *duǒ* 'lit. flower (for flowers, clouds, etc.)'
c. conjunctions: *hé* 'lit. union (general purpose)', *yǔ* 'lit. accompany (formal, literary)'

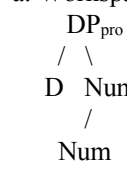
These items can all be analyzed as a functional category supported by a root (e.g., [$\text{Voice}_{\text{PASS}}$ √BÈI]). The categorial part determines their syntactic functionality, while the root part determines their extragrammatical effects and thereby conditions their real-life usage (e.g., the "and" in *Harry Potter and ...* is *yǔ* instead of *hé*).

Returning to noncanonical pronouns, the same analysis applies. Since we have established that the syntactic behavior of noncanonical pronouns is the same as that of default pronouns, we treat them as root-supported $D_{pro}$ items. However, given the elaborate DP structures in §4.1, we next clarify three technical details to show how our analysis fits into the big picture of pronominal syntax.

### *4.3 Deriving noncanonical pronouns*

First, following Nunes and Uriagereka (2000), Johnson (2003), and especially Zwart (2007 et seq.), we assume syntactic derivation to be multilayered. That is, derivational products of an earlier cycle/workspace can be used in a subsequent cycle, probably in an "atomized" fashion (Fowlie 2013). This is an inevitable state of affairs if we look closely at the assembling of syntactic trees. At each step of Merge, the object that is newly selected from the lexical (sub)array into a workspace by definition has not undergone any Merge step in that workspace (so it is a "minimal" category in a relative sense). While this is natural in the merger of a primitive category with an existing phrase, a question arises in the merger of a specifier or adjunct with a phrase: Where has the specifier/adjunct (which is also phrasal by definition) been derived? The only logical possibility is that it has been derived in another workspace before being selected and merged in the current workspace. The upshot is that "layered derivation" (Zwart's term) must be a standard mechanism in minimalist syntax. We contend that this is also what happens in the derivation of noncanonical pronouns.

(31) a. Workspace I

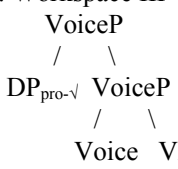



(atomize $DP_{pro}$)

b. Workspace II


(support $DP_{pro}$ with a root)
(atomize the root-supported $DP_{pro}$)

c. Workspace III




(merge the root-supported $DP_{pro}$ as a specifier)

The above derivation involves three consecutive layers, each defined by a workspace. In (31a), a default pronominal phrase is built up in Workspace I. We use the DP-NumP-NP structure as an example, but as mentioned in §4.1, any workable DP structure is fine for current purposes. In (31b), the assembled and atomized pronominal DP is selected into Workspace II and merged with a root. This is the categorization step. As a result, the root is assigned the category D and the D category is associated with the idiosyncratic content in the root. Finally, in (31c) the root-supported pronominal DP is selected into Workspace III and merged as a specifier. The particular scenario here is Spec-VoiceP; namely, the subject of a transitive verb.

Second, the derivation in (31) does not hinge on the concrete content in the supporting root, since that is syntactically inert anyway. This means that miscellaneous lexical materials can be (re)used at the root slot, as long as it is properly lexicalized, with its original formal features (if any) reanalyzed as lexical features. As such, not only simple roots like √ZHÈN but also derived roots like √ĀI-JIĀ can be used to support $DP_{pro}$. The key prediction here is that this recategorized *āijiā* can only be used as a pronoun but not as a common R-expression anymore, even though its original meaning clearly *is* R-expressional (i.e., "mourner"). As we observed in §2, this is generally true for noncanonical pronouns. Third, the tree in (31a) is just that of a default pronoun, which in principle can have its own overt form. For instance, if the pronoun is 1SG, then it in principle can be spelled out as the default *wǒ* in Chinese. Yet this never happens with noncanonical pronouns. That is, noncanonical pronouns do not allow appositive default pronouns, which sharply contrasts them with imposters, as in (32).[26]

(32) a. *Lǎoshī (wǒ) kuài yào shīqù wǒ de nàixìng le.* [Mandarin, imposter]
teacher.1SG 1SG quick going.to lose 1SG POSS patience CRS
"Teacher$_{1SG}$ (I) is going to lose my patience." (adapted from Want 2014:185)

b. *Zhèn (*wǒ) kuài yào shīqù wǒ de nàixìng le.* [noncanonical pronoun]
zhèn.1SG 1SG quick going.to lose 1SG POSS patience CRS
"Zhèn$_{1SG}$ (*I) is going to lose my patience."

Intuitively, *zhèn* is just *wǒ* with an alternative pronunciation and some idiosyncratic effects, so saying *zhèn wǒ* is like saying "I I" in English, which is clearly ungrammatical. How can the syntactic derivation in (31) bear this out, though? The observation, in conditional terms, is the following:

(33) i. If a functional category is root-supported, it assumes the root's exponent (possibly distorted).
ii. If a functional category is non-root-supported, it assumes its default exponent (if any).

There are two ways to explain (33). One way is to view it as a concomitant of the categorization procedure. Recall that the classic case of categorization in distributed morphology is that of content words. In this case the uncategorized root has no fixed pronunciation and only gets one when it is assigned a category. For instance, the English root √PERMIT may be pronounced as /pɚˈmɪt/ or /ˈpɚˌmɪt/ depending on its category. This is even more evident in languages like Hebrew, where uncategorized roots cannot be vocalized at all (e.g., √K-T-B 'related to writing', √L-M-D 'related to learning'). The

[26] While imposters may co-occur with appositive default pronouns in Chinese, this is impossible in Vietnamese.

theory for this in distributed morphology is that each step of categorization in syntax corresponds to a step of retrieving stored phonological/semantic information at the interfaces. And crucially, it is the [categorizer root] unit as a whole, not its subparts, that gets assigned that information. As such, it is plausible that each noncanonical pronoun has its own lexical entry, and when a root-supported $DP_{pro}$ like (31b) is interpreted, it is assigned the stored pronunciation, which in most cases is the same as that of the recycled lexical material (again a manifestation of MMM) but may also involve certain distortion as in *lúnjiā*.

Another way to explain (33) is to invoke Kiparsky's (1973) elsewhere principle, which basically says that when a more general and a more specific rule are adjacent, the more specific rule applies. For instance, the past tense form of *go* is *went* instead of **goed* because the irregular rule is more specific. Neeleman and Szendrői (2007) use the elsewhere principle to explain radical pro drop. We leave out the details of their application for space limitations but merely cite a well-known implication of the principle that they list:

(34) All else being equal, the phonological realization of syntactic structures favors spell-out of a category C over spell-out of the categories contained in C. (Neeleman and Szendrői 2007:685)

This immediately explains why a root-supported functional category gets pronounced differently from a non-root-supported one—because in a tree like (31b) the spell-out rule that targets [$DP_{pro}$ √] as a whole blocks the rule that targets $DP_{pro}$ itself. We can reformulate (33) as follows to better reflect the elsewhere principle.

(35) i. If a functional category is root-supported, it assumes the root's exponent (possibly distorted).
ii. Elsewhere it assumes its default exponent (if any).

In any event, the different pronunciations of default pronouns and noncanonical pronouns, in spite of their partially identical underlying structures, well conform to independently motivated rules in generative syntax.

### *4.4 Crosslinguistic availability*

We mentioned in §1 that noncanonical pronouns have a much more limited crosslinguistic distribution than imposters. Our foregoing analysis may explain why this should be the case. Song (2019:136–137) points out that root-supported heads are "analytic heads" because they have a very low category/morpheme-per-word ratio (respectively 1:1 and 2:1), where a word is understood as a morphophonologically freestanding unit. And Chinese-style semifunctional items (like those in (30)) are furthermore "analytic heads par excellence" because their category-per-word ratio and morpheme-per-word ratio are both 1:1, where a morpheme is understood in the traditional sense as a minimal sound-meaning pair. As such, by analyzing noncanonical pronouns in generalized root syntax, we automatically get the following prediction:

(36) Noncanonical pronouns are more common in highly analytic languages.

This may explain why we can easily find noncanonical pronouns in Vietnamese and Chinese but not in familiar European languages—both Vietnamese and Chinese are highly analytic languages, which have the right grammatical setting for root-supported heads to stably exist. Note that we are not claiming that high analyticity is solely defined by root support. In fact, if there is a high analyticity parameter at all, that is very likely to be codefined by a cluster of smaller parameters or grammatical settings (see Huang 2015 for a discussion). Here we are merely stating that a highly analytic language has a root-support-friendly setting. For instance, among others it may have scarce head movement, which means it has more standalone words than affixes, and root support just provides a convenient way to create standalone grammatical words. Of course, we need to check more highly analytic languages to verify (36), which we leave to future research.

## 5 Conclusion

In this paper, we studied a new type of pronominal item emerging on the Internet in Vietnamese and Chinese. First, we demonstrated that pronominal items of this new type, which we call noncanonical pronouns, are a separate category from both textbook default pronouns and imposters (§1). Then, we illustrated their real-life usage in detail (§2). Our investigation showed that noncanonical pronouns in the two Asian languages are similar not only in syntactic behavior but also in lexical sources and evolution pathways. Syntactically, both Vietnamese and Chinese noncanonical pronouns behave like default pronouns except that they have various extragrammatical effects, mostly reflecting the speaker's mind-set or personality. Lexically, both languages subsume three subtypes of noncanonical pronouns based on their evolution pathways: revived ancient terms, dialectal terms, and creative online coinages. After presenting the empirical facts, we briefly discussed the taxonomy of pronominal items and specified why we prefer a syntactically based taxonomy (§3).

In the theoretical part of the paper, we analyzed noncanonical pronouns in the theory of generalized root syntax (§4), which is an extension of the root theory in distributed morphology (a branch of the minimalist program). Specifically, we analyzed noncanonical pronouns in the schema [$DP_{pro}$ √], where $DP_{pro}$ is a separately derived and atomized default pronoun structure and √ is a purely lexical root supporting that structure. The root part may be constituted by terms (re)lexicalized from the three sources above. After the (re)categorization, the root-supported structure $DP_{pro\text{-}√}$ is selected into the main workspace and merged onto the main tree, where it behaves like a default pronoun in syntax but triggers idiosyncratic phonological and semantic properties at the interfaces, just as we have observed in Vietnamese and Chinese. Alongside our analysis, we clarified a number of theoretical and technical issues, such as how to use previous theories of pronominal syntax in a well-founded way, why cross-workspace or layered derivation must be allowed in generative syntax, and why default and noncanonical pronouns with the same pronominal structure can have totally different overt forms. Finally, we also tentatively explained why noncanonical pronouns have limited crosslinguistic availability in terms of the correlation between root support and high analyticity.

Due to limited scope, we have had to leave some interesting questions to future research, including but not limited to the acquisitional order predicted by our taxonomy in §3 and the availability of noncanonical pronouns in other highly analytic languages as predicted in §4. We also observed a quasi-complementary contrast between Vietnamese and Chinese in §2 concerning the distribution of noncanonical pronouns and imposters. So, as a future plan we would also like to further compare Vietnamese and Chinese imposters and look into questions such as why imposters are used more freely in Vietnamese (footnote 4), why they may co-occur with appositive default pronouns in Chinese but not in Vietnamese (footnote 26), and so on.

## Abbreviations

1/2/3 = first/second/third-person
AFFECT = affectionate
ARG = argument
CLF = classifier
COP = copula
CRS = currently relevant state
DEM = demonstrative
DET = determiner
DIM = diminutive
DISP = disposal
DM = discourse marker
EMP = emphatic
EXCL = exclamative
EXP = experiential
FEM = feminine
FUT = future
HON = honorific

IMPERF = imperfective
INTS = intensifier
MASC = masculine
N = noun
NEG = negation
NMLZ = nominalizer
Num = number
OFF = offensive
PL = plural
POSS = possessive
PRF = perfective
PROG = progressive
PRT = particle
Q = question marker
REL = relative clause marker
RES = resultative
SG = singular
TEL = telic marker
TRANS = transitive

## References

Abney, Steven. 1987. The English noun phrase in its sentential aspect. MIT dissertation.

Alves, Mark. 2007. Categories of Sino-Vietnamese grammatical vocabulary. *Mon-Khmer Studies* 37:217–229.

Alves, Mark. 2017. Chinese influence on Vietnamese pronouns and terms of address and reference. Presentation at *NACCL 29*, June 16–18.

Acquaviva, Paolo. 2019. Categorization as noun construction: Gender, number, and entity types. *Gender and noun classification*, ed. by Eric Mathieu, Miriam Dali and Gita Zareikar, 41–63. Oxford: Oxford University Press.

Aldridge, Edith. 2021. The syntax of pronoun fronting in Late Archaic Chinese negated clauses. *Journal of East Asian Linguistics* 30.1:39–79.

Biberauer, Theresa. 2011. In defence of lexico-centric parametric variation: Two 3rd factor-constrained case studies. Paper presented at Workshop on Formal Grammar and Syntactic Variation, Oct 28, Madrid.

Biberauer, Theresa. 2017. Factors 2 and 3: A principled approach. *Cambridge occasional papers in linguistics*, vol. 10, ed. by Chechens Song and James Baker, 38–65.

Borer, Hagit. 2005. *In name only*. *Structuring sense*, vol. 1. Oxford: Oxford University Press.

Cao, Xuan Hai. 2014. Đại từ nhân xưng đích thực trong tiếng Thanh Hoá [De facto personal pronouns in the dialect of Thanh Hoa]. *Ngôn ngữ và Đời sống* 9.27:17–22.

Cardinaletti, Anna; and Michal Starke. 1999. The typology of structural deficiency. *Clitics in the languages of Europe*, ed. by Henk van Riemsdijk, 145–233. Berlin: Mouton de Gruyter.

Chen, Cuizhu. 2009. Hanyu rencheng daici kaolun [An examination of Chinese pronouns]. Central China Normal University dissertation.

Chomsky, Noam. 1981. *Lectures on government and binding*. Berlin: Mouton de Gruyter.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge MA: MIT Press.

Chomsky, Noam. 2013. *Problems of projection*. *Lingua* 130:33–49.

Collins, Chris. (ed.) 2014. *Cross-linguistic studies of imposters and pronominal agreement*. Oxford: Oxford University Press.

Collins, Chris; and Paul Postal. 2012. *Imposters: A study of pronominal agreement*. Cambridge MA: MIT Press.

Collins, Chris; and Edward Stabler. 2016. A formalization of minimalist syntax. *Syntax* 19.1:43–78.

Déchaine, Rose-Marie; and Martina Wiltschko. 2002. Decomposing pronouns. *Linguistic Inquiry*, 33.3:409–442.

Deng, Da. 2021. Wangluo chengwei qin de yanjiu [A study of the online appellation qin]. Xiangtan University MA thesis.

Feng, Shengli. 2010. Lun yuti de jizhi ji qi yufa shuxing [On mechanisms of register system and its grammatical property]. *Studies of the Chinese Language* 5:400–412.

Fowlie, Meaghan. 2013. Multiple multiple spellout. *Challenges to linearization*, ed. by Theresa Biberauer, 129–169. Berlin: Mouton de Gruyter.

Galloway, Brent. 1993. *A grammar of Upriver Halkomelem*. Berkeley: University of California Press.

Halle, Morris; and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. *Essays in linguistics in honor of Sylvain Bromberger*, ed. by Kenneth Hale and Samuel Jay Keyser, 111–176. Cambridge MA: MIT Press.

Harbour, Daniel. 2016. *Impossible persons*. Cambridge MA: MIT Press.

Hoang, Thi Chau. 1989. *Tiếng Việt trên các miền đất nước: phương ngữ học* [Vietnamese in the various areas of the motherland: a dialectological study]. Hà Nội: NXB Khoa học Xã hội.

Huang, C-T. James. 2015. On syntactic analyticity and parametric theory. In *Chinese syntax in a cross-linguistic perspective*, ed. by Audrey Li, Andrew Simpson and Wei-Tien Tsai, 1–48. Oxford: Oxford University Press.

Johnson, Kyle. 2003. Towards an etiology of adjunct islands. *Nordlyd* 31.1:187–215.

Kiparsky, Paul. 1973. "Elsewhere" in phonology. *A festschrift for Morris Halle*, ed. by Stephen Anderson and Paul Kiparsky, 93–106. New York: Holt, Rinehart and Winston.

Liu, Yang. 2012. Taobaoti zhong qin de yanjiu [A study of qin in Taobaospeak]. Heibei Normal University MA thesis.

Neeleman, Ad; and Kriszta Szendrői. 2007. Radical pro drop and the morphology of pronouns. *Linguistic Inquiry*, 38.4:671–714.

Neeleman, Ad; and Fred Weerman. 1999. *Flexible syntax: A theory of case and arguments*. Dordrecht: Kluwer.

Nunes, Jairo; and Juan Uriagereka. 2000. Cyclicity and extraction domains. *Syntax* 3.1:20–43.

Pollock, Jean-Yves. 1989. Verb movement, universal grammar, and the structure of IP. *Linguistic Inquiry* 20.3:365–424.

Postal, Paul. 1966. On so-called pronouns in English. *Modern studies in English*, ed. by David Reibel and Sanford Schane, 201–223. Englewood Cliffs, N.J.: Prentice-Hall.

Ritter, Elizabeth; and Martina Wiltschko. 2019. Nominal speech act structure: Evidence from the structural deficiency of impersonal pronouns. *Canadian Journal of Linguistics*, 64.4:709–729.

Ritter, Elizabeth. 1995. On the syntactic category of pronouns and agreement. *Natural Language & Linguistic Theory* 13:405–443.

Rizzi, Luigi. 1997. The fine structure of the left periphery. *Elements of grammar: A handbook of generative syntax*, ed. by Liliane Haegeman, 281–337. Dortrecht: Kluwer.

Roberts, Ian. 2019. *Parameter hierarchies and universal grammar*. Oxford: Oxford University Press.

Song, Chenchen. 2019. On the formal flexibility of syntactic categories. University of Cambridge dissertation.

Wang, Chyan-An. 2014. Mandarin pseudo-imposters. *Cross-linguistic studies of imposters and pronominal agreement*, ed. by Chris Collins, 173–195. Oxford: Oxford University Press.

Weerman, Fred; and Jacqueline Evers-Vermeul. 2002. Pronouns and case. *Lingua* 112:301–338.

Xia, Xianpei. 2018. Nüzi zicheng chenqie de yuanliu [On the origin of the female self-appellation chenqie]. *Journal of Changsha University of Science and Technology (Social Science)*, 33.5:118–122.

Zwart, Jan-Wouter. 2007. Layered derivations. ACTL lecture. London: University College London.