Chapter 6

Borrowed affixes and morphological productivity: A case study of two Maltese nominalisations

Albert Gatt
Institute of Linguistics and Language Technology, University of Malta

Ray Fabri
Institute of Linguistics and Language Technology, University of Malta

Among the derivational processes that have been adopted into Maltese based on the Romance model, there are processes to derive nouns from verbs which are relatively recent developments. Examples include the use of the suffix *-Vr* (e.g., *spara/sparar* 'shoot'/'(the) shooting'), and the use of *-(z)zjoni* (e.g., *spjega / spjegazzjoni* 'explain'/'explanantion'). This paper discusses these processes in the context of Maltese derivation in general. After a brief theoretical exposition and an overview of Maltes derivation, we present a corpus-based analysis of the productivity of *-Vr* and *-(z)zjoni* derivations, followed by an analysis of the evidence for indirect borrowing in these two cases, based on the work of Seifart (2015). We show that, while there is evidence that both are productive, the statistical evidence suggests that *-Vr* processes are more likely to result in novel forms. By the same token, *-Vr* nominalisations more clearly represent cases of indirect borrowing, as evidenced by the greater number of types which have corresponding simplex forms, and by the greater probability that the simplex forms are more frequent than the nominalisations.

1 Introduction

The morpho-phonological system of Modern Maltese is the result of intensive contact involving an Arabic stratum, a Romance (Sicilian, Italian) superstratum

Albert Gatt & Ray Fabri. Borrowed affixes and morphological productivity: A case study of two Maltese nominalisations. In Patrizia Paggio & Albert Gatt (eds.), *The languages of Malta*, 143–169. Berlin: Language Science Press.
DOI:10.5281/zenodo.1181793

and an English adstratum (Brincat 2011; Mifsud 1995b). The result is a hybrid morphological system incorporating both root-based and stem-based forms (see, among others, Drewes 1994; Mifsud 1995b; Fabri 2010; Spagnol 2013).

The extent to which the Semitic or Romance components predominate in the Maltese lexicon is something of an open question. For example, Brincat (2011) suggests that the Romance component accounts for some 52% of the lexicon, based on counts obtained from a standard dictionary. By contrast, Comrie & Spagnol (2016) restrict their counts to a small sample of lexical items and find that Romance etymology accounts for some 30%.

Whatever the actual predominance of Romance versus Semitic, the hybridity of the Maltese morphological system raises a number of empirical questions which have broad theoretical significance. Among these is the question of productivity. The term 'productivity' is here operationally defined as referring to the extent to which morphological processes are used by speakers to generate novel forms. This criterion – i.e. the use of a process to generate novel forms – is widely adopted in many discussions of morphological productivity (including Aronoff & Schvaneveldt 1978; Cutler 1980; Aronoff & Anshen 1998; Bauer 2001; Dressler 2003; and Plag 2004, among others.)

In earlier work on Maltese (for example by Mifsud 1995b; Hoberman & Aronoff 2003), it has sometimes been argued that the Semitic/Arabic derivational component is relatively unproductive, with novel word forms being largely created through stem-based processes ultimately arising from Romance. Indeed, some recent work by Saade (2016) has suggested that quantitative measures of the productivity of a subset of Romance-derived derivational affixes in Maltese yield a ranking comparable to that found for their Italian cognates in a earlier work by Gaeta & Ricca (2006), although the absolute productivity values found are lower for Maltese processes compared to their Italian counterparts.

These views and findings motivate the decision in the present study to focus primarily on comparison of two non-Arabic deverbal suffixes, namely *-(z)zjoni* and *-Vr* (where *V* is either /a/ or /i/), as a test case. Thus, certain derivational processes of Arabic origin fall outside the scope of the present paper, though we discuss them briefly in order to situate our study within the broader context of the Maltese morphological system.

We conduct a corpus-based analysis of *-Vr* and *-(z)zjoni* to address two related issues: (a) to what extent these affixes are productive, based on statistical criteria formulated by Baayen (2009); and (b) whether there is evidence for their status as directly or indirectly borrowed affixes, in the sense discussed by Seifart (2015). Individually, these two questions shed light on the nature of the morphological

system in Maltese and the extent to which empirical evidence justifies claims that certain processes are in greater use than others in modern Maltese. However, the questions of productivity and of whether an affix is in/directly borrowed are also complementary in another important sense. Following Seifart (2015) we view indirect borrowing as a process involving the importation of complex forms having an affix from a source language, followed by the use of that affix on novel forms, thus implying parsing or decomposition. This implies that productivity is part and parcel of the definition of indirect borrowing. As Hay & Baayen (2001) argue, the extent to which a lexical item can be parsed in perception (for example, into a stem and its affixes) can help in predictions of the productivity of the processes giving rise to that item in the first place. By the same token, the ability to place an affix (or morphological process) somewhere along the continuum between direct and indirect borrowing complements the statistical evidence for productivity that can be derived from corpora along the line suggested by Baayen (2009), among others.

This chapter is structured as follows. We first briefly introduce the theoretical framework within which we view derivational morphological processes. Next (§3), we discuss derivation in Maltese, with particular reference to nominalisation and stem-based processes. §4 gives a descriptive overview of the two affixes under consideration, with remarks concerning their status as directly or indirectly borrowed affixes. §5 describes a corpus-based investigation of productivity of *-Vr* and *-(z)zjoni* forms, followed by an empirical investigation of the evidence in favour of indirect borrowing in the two cases. §6 concludes with some remarks on possible future research directions.

2 Some theoretical preliminaries

In the present study, the term 'derivation' is not meant to imply a process that has a direction, as in 'X is derived from Y' (as opposed to 'Y is derived from X'). In other words, we do not adopt a procedural approach to derivational morphology, which would require a commitment as to the precedence of certain forms from which others are derived. Even for the linguist, the decision as to which lexemes in a pair has (diachronic) priority is often a matter of approximate reasoning (as Ellul 2016, for example, shows in connection with Maltese deverbal nominalisations).

Rather, we propose to think of derivation in terms of links between forms that have a formal and semantic relationship, that is, between words/lexemes. The term 'derivation' is used here simply to distinguish one type of lexical relation from other relations, in particular from inflection, which is a relation between

grammatical word forms instantiating a lexeme, as opposed to a relation between lexemes. In the following, we use a double arrow (↔) to express the idea that in X ↔ Y, X and Y are derivationally related to each other. Thus, for example, in English, REVOLT ↔ REVOLUTION means that *revolt* and *revolution* are two lexemes that have both a formal and a semantic relationship.

One way to conceive of such derivational relationships in a non-procedural fashion is with reference to what we call *derivational families*, i.e., words/lexemes that are related to each other through a common, shared base. Cases that involve allomorphic variation of the base (including stem allomorphy) are also included within a derivational family. To take an example, the base form KOMPO 'compose/make up' relates the following forms to each other: *(ik)kompo(n)-a* 'compose', *kompo(n)-iment* 'essay', *kompo(n)-enti* 'component', *kompo(ż)-izzjoni* 'composition' and *kompo(ż)-itur* 'composer'.

The assumption is that the base in all of these forms somehow expresses some common underlying, basic meaning, or at least serves to index a cluster of related basic meanings. To the native speaker, the relationships would be intuitively obvious for *kompożizzjoni* 'composition' in relation to *kompożitur* 'composer', though perhaps less obvious for *komponent* 'component' relative to *komponiment* 'essay'. At the same time, the relationship between the latter two, while semantically more opaque than in the case of KOMPOŻIZZJONI ↔ KOMPOŻITUR, is nevertheless quite clear from a formal perspective (i.e. the shared stem is intuitively obvious to the native speaker). This suggests that there may be a disjuncture between the semantic and formal links, so that one could conceive of a combination of at least two criteria governing the intuition of a relationship, one based on form and another on meaning, where the strength of the relation between two forms can vary as a function of the strength of the formal and semantic relations. Thus, a more tenuous link would be perceived between two forms if, say, they share a base but the meaning deviates considerably (see Bybee 1995 for a network model of morphology developed along these lines).

A gradient of associative strength among related forms would also be able to account for derivational relationships among forms which are related by a common root via processes of Arabic origin, as in the case of $\sqrt{b\text{-}j\text{-}d}$ in (1) and $\sqrt{ħ\text{-}d\text{-}m}$ in (2).

(1) (Second verbal form (CVCCVC), from *abjad*, root $\sqrt{b\text{-}j\text{-}d}$;)
abjad ↔ *bajjad*
white ↔ paint/whitewash
to whitewash

(2) (Deverbal noun from root $\sqrt{\text{ħ-d-m}}$;)
ħadem ↔ *ħaddiem*
work ↔ worker
worker

These might be presumed to constitute families in which the root indexes a cluster of basic meanings, though there is considerable semantic variation among related forms. This need not imply that the processes involved are productive (Mifsud 1995b; Hoberman & Aronoff 2003), or that all relationships among lexemes sharing a root are equally transparent – indeed, as we have seen, some arguments to the contrary have been advanced. It does, however, imply that the root itself has some psychological reality for the native speaker. It is worth noting that some psycholinguistic evidence does point towards a role for the root as an index for lexical retrieval by native Maltese speakers (see e.g. Twist 2006; Ussishkin et al. 2015).

Another frequently observed phenomenon is that derivational families (unlike inflectional families, generally) display gaps, that is, not all the theoretically or potentially possible forms actually occur in everyday use. To take an example, the Korpus Malti, a corpus of Maltese (described in Gatt & Čéplö 2013 and introduced more fully in §5 below), yields examples of a number of forms for the base DIMOSTR- 'demonstrate', including *dimostrazzjoni* 'demonstration', *dimostrant* 'demonstator', *(i)ddimostra* 'demonstrate', and even one occurrence of *dimostratur* 'demonstrator' (the more frequently attested form being *dimostrant*). The corpus does not, however, attest to the use of potential formations such as *#dimostrist* or *#dimostrament*, though the suffixes *-ist* and *-ment* are productively used in Maltese. The established family networks provide the potential for new creations, but many factors play a role in determining which are actually formed and used, including phonological restrictions on the base, blocking through an already available form, and other constraints which have been discussed extensively in the theoretical literature (e.g. Spencer 1991; Spencer & Zwicky 1998; Aronoff & Fudeman 2011; and Haspelmath & Sims 2010, among others.)

3 Derivational morphology in Maltese

As noted in the introductory section, at different phases of its history, Maltese borrowed lexically from a number of languages. Its early sources were mainly Sicilian, Tuscan, and Modern Italian; in recent times, English has become an additional source (for discussion of these various influences see Mifsud 1995b; Borg &

Azzopardi-Alexander 1997; Fabri et al. 2013; Spagnol 2013; and Brincat & Mifsud 2016). As a result, Maltese displays a great deal of lexical and morphological variety, and derivation also reflects this rich historical background, displaying both non-concatenative (templatic, root-based) forms (exemplified in 1 and 2 above), which are generally older forms historically going back to Arabic, and concatenative (affixal, stem-based) forms, which are generally historically of non-Arabic origin, i.e., Sicilian, Italian or English.

In this section, we first give a brief overview of the types of derivational processes available, before turning to a consideration of the status of stem-based derivational processes, in anticipation of the study presented in §5.

3.1 Verbal derivation

The historically older derived verbal forms are based on the conjugation system typical of Arabic, often referred to with the term *binyanim* from Hebrew, and known as *forom* 'forms' in Maltese. Traditional descriptive grammars list 10 derivational verbal forms, though the vast majority of Maltese verbs do not conjugate in all forms (in fact, the majority have only between two and three forms, as shown by Spagnol 2013) and at least one form – form IV – has only a single attested entry and thus cannot really be considered a derivational form in modern Maltese. The roots are generally assumed to be triliteral, as in the case of $\sqrt{d\text{-}ħ\text{-}l}$ 'enter'; or quadriliteral, as in the case of $\sqrt{ħ\text{-}r\text{-}b\text{-}t}$ 'spoil/ruin'. The derived forms are characterised either by changes in the CV template (i.e., non-concatenative processes), by affixation, or both. Table 1 displays a few examples in addition to those given in (1) and (2) above.

Table 1: Examples of root-based verbal derivations.

Root	Derived form	Form #	Gloss
$\sqrt{d\text{-}ħ\text{-}l}$	daħħal	II	'let in'
$\sqrt{f\text{-}h\text{-}m}$	fiehem	III	'explain'
$\sqrt{k\text{-}s\text{-}r}$	tkisser	V	'get broken'
$\sqrt{d\text{-}ħ\text{-}l}$	ndaħal	VII	'interfere'

The productivity of these derivational forms in modern Maltese, which evince a lot of gaps and are often semantically idiosyncratic, is a matter of discussion (see, e.g., Mifsud 1995b and Hoberman & Aronoff 2003). There appears to be general agreement, based mainly on intuition, that these forms are fossilised and not

generally productive, that is, new formations in the templatic system are rare to non-existent. A separate issue, which we noted in our theoretical outline in §2, is whether the root has any psychological reality. Evidence from studies of lexical access has suggested that this is the case (Twist 2006; Ussishkin et al. 2015).

Note that, although new verbal forms are not being created within the templatic system, new verbs, especially from English, are being created in Maltese (these verbs are referred to as Type D verbs by Mifsud 1995a). These are created on the pattern of a particular declensional class of verbs, namely, verbs with a final weak consonant, that is, *j* or *w*. The new forms are characterised by the suffix *-ja* attached to a borrowed base form, which is either verbal or nominal in origin, to produce a verbal stem for inflection. Often this process is also accompanied by gemination of the initial consonant, which then requires *i*-epenthesis for syllabification, as in *immoniterja* 'monitor', from English *monitor*; and *iċċekkja* 'check', from English *check*. This process appears to be highly productive, with new verbs continually being produced according to this pattern, as shown by Mifsud (1995a). These verbs, in turn, become candidates for deverbal derivation in forms such as *iċċekkjar* 'checking', formed using *-Vr* , which will be discussed below.

3.2 Nominal derivation

Derived nominals (nouns and adjectives) also consist of formations that display both concatenative (i.e., affixal) and non-concatenative (i.e., templatic) patterns. Table 2a shows a few examples of noun patterns that are derivationally related to other forms via templatic processes. Table 2b gives some examples of nominal derivations which arise from affix-based processes.

There are indications that, just as in the case of verbal derivation based on templatic patterns, templatic nominal derivation is not productive anymore.

3.3 The status of stem-based derivational processes

Many of the stem-based derivational processes outlined above raise the question whether they involve 'real' affixes. Clearly, whether or not they are productive is an important consideration here. Justification for treating such affixes as productive morphemes generally comes from cases of local formations which do not have cognates in a source language, since this means that they could not have been absorbed whole but must have been created locally.

Obvious examples of local creations are derived forms which have a lexical base from one language source but which make use of a derivative feature (affix-

Table 2: Examples of templatic (root-based) and affix-based nominal derivation patterns.

Template	Root	Root Gloss	Example	Gloss
CVCC/a	$\sqrt{s\text{-}r\text{-}q}$	steal	serq/a	(a) theft
t-VCCiiC/a	$\sqrt{\hbar\text{-}w\text{-}d}$	mix up	taħwid/a	(a) mix-up
CCVVC	$\sqrt{\dot{z}\text{-}f\text{-}n}$	dance	żfin	(the) dancing
CVC_iC_iVVC	$\sqrt{\hbar\text{-}d\text{-}m}$	work	ħaddiem	worker

(a) Template-based patterns

Affix	Base	Base Gloss	Example	Gloss
-ment	aġġorna	to update	aġġornament	the/an update
-tur/a	ċċekkja	to check	ċekkjatur/a	checker/a check
-ist	arti	art	artist/a	artist
-vġġ	arpa	harp	arpeġġ	arpeggio
-vr	spara	to shoot	sparar	the/a shooting
-(z)zjoni	kkonserva	to conserve	konservazzjoni	conservation

(b) Affix-based patterns

ation, templatic arrangement) from a different language source. The examples in (3) show the well-known case of the Italian origin suffix *-ata* being attached to stems of words of Arabic origin to create new lexemes.

(3) (Suffix *-ata* applied to stems of Arabic origin;)
fenek ↔ *fenkata*
rabbit ↔ rabbit meal
xemx ↔ *xemxata*
sun ↔ sunstroke

At first blush, this suggests that such affixes have made their incursion into Maltese through what Seifart (2015) calls *indirect borrowing*, which Seifart places at one end of a continuum, at the other end of which is *direct borrowing*. In the latter case, ‘an affix is recognized by speakers of the recipient language …and used on native stems as soon as it is borrowed, with no intermediate phase of

occurrence in complex loanwords only' (p. 512). By contrast, the paradigm case of indirect borrowing occurs where a number of lexical items with a particular affix are first borrowed into the target language, with the affix gradually coming into productive use on native stems following a process of analysis of the borrowed items. Note that this characterisation of direct versus indirect borrowing is diachronic in flavour. However, Seifart also suggests a number of criteria for identifying an indirectly borrowed affix in a language at a given stage of development. We turn to these in §4.2 below, where we discuss the question of whether the two nominalising affixes under discussion are best thought of as examples of direct or indirect borrowing.

4 *-Vr* and *-(z)zjoni* nominalisations

Following the overview above, we now turn our attention to a case study involving two nominalisation suffixes in Maltese: *-Vr* and *-(z)zjoni* . Before we present the results of a quantitative investigation, we give a descriptive outline.

4.1 Descriptive outline

-Vr is usually classified as *-ār*, with long /a/, and traced back to the Italian infinitive ending *-āre*, as in *amare* 'love' (Mifsud 1995a: 249). Indeed, in Italian the infinitive form can function as a noun, as shown by the use of *dire* 'say' and *fare* 'do' in the following proverb: *tra il dire e il fare c'è di mezzo il mare* (literally: 'there is an ocean between the said and the done'). However, while the *-are* ending (equivalently *-ere* and *-ire*) in Italian is not specifically a nominaliser, but marks the verb as infinitive, though it can then be used as a (verbal) noun, *-Vr* in Maltese is specifically and exclusively a nominaliser. Indeed, Maltese, like other Semitic languages, does not have a morphological infinitive.

The *-Vr* ending can be found both with Italian stems, as in *issorveljar* '(the) overseeing' from Italian *sorvegliare* 'oversee', and with English stems, as in *ibbrejkjar* '(the) braking' and *ipparkjar* 'the parking', from English *brake* and *park*, respectively. Given that, in these cases, the Maltese verbal form ends in short /a/ (e.g. *ipparkja* and *ibbrejkja*), the assumption is usually that the *-Vr* nominal is related to a verbal stem which already displays the /a/. There are however a few forms which display an *-īr* in place of *-ār*. Examples are *aġir* 'action', *servir* 'serving', *avvertir* 'warning', *riferir* 'referral', *esegwir* 'execution (of an action)' and *distribwir* 'distribution' (see Camilleri 1993 for a complete listing). To be sure, these are far less frequent than the forms involving /a/.

Like *-ār* forms, *-īr* forms are assumed to be related to a verbal stem ending in /a/, as in *irrefera* 'refer' or *esegwixxa* 'execute'. Interestingly, although these forms end in *-a* in the perfect third person masculine singular, in the imperfect singular they end in /i/, thus, *tirreferi* 'you refer', *teseġwixxi* 'you execute/she executes'. Moreover, these verbs are historically derived from verbs which in Italian end in *-ire* (*riferire*, *eseguire*). On being integrated into the Maltese inflectional system, they came to be conjugated on the pattern of a set of verbs of Arabic origin, such as *ħeba* 'hide' and *qela* 'fry', which end in /i/ in the imperfect singular (cf. *naħbi* 'I hide', *taħbi* 'you hide/she hides', *jaħbi* 'he hides'). Arguably, the *-i* in these cases can be taken as an inflectional suffix for the imperfect singular, as opposed to *-u* for plural (cf. *tirreferu* 'you (plural) refer', *taħbu* 'you (plural) hide'). In any case, these verbs contrast with the more common stem ending in *-a*, such as *tissorvelja* 'you oversee' and *tibbrejkja* 'you brake'.

Although traditionally the third person masculine singular perfect form (called *il-mamma* in Maltese pedagogical grammars) is taken as the citation form and often as the base form, speakers more naturally produce the second person singular as citation form when asked to give a Maltese equivalent for a foreign verb. This might be taken as an indication that, to the native speaker, the intuitive base form is indeed the second person singular, with the stem ending in /i/ or /a/ explaining the difference between forms such as *ibbrejkjar* 'to brake' (from second person *tibbrejkja*) and *avvertir* 'to warn' (from second person *tavverti*). This is why we use *-Vr* rather than *-ār* to indicate the relevant morph. Nevertheless, as noted by Camilleri (1993), *-īr* forms are comparatively rare.

The suffix *-(z)zjoni* comes from Italian *-zione* (compare: *ġeneralizzazzjoni* 'generalisation', from Italian *generalizzazione*) and has probably been 'strengthened' by English *-ation*. Thus, for example, the Maltese forms *afforestazzjoni* 'afforestation' and *aġġudikazzjoni* 'adjudication' do not have obvious cognates in Italian but they do have English equivalents in *-ation*. Here, too, there are candidates for allomorphic variants of the suffix, whose status is however unclear. Relevant examples are *manutenzjoni* 'maintenance', *intenzjoni* 'intention', and *prekawzjoni* 'precaution', all of which have a singulative /z/, rather than a geminate. (Note that this is not an orthographical but a phonological effect.) The former are preceded by a stem-final consonant, the latter by a stem-final vowel; cf. *manuten-zjoni* vs. *assoċja-zzjoni*. For this reason, we characterise the suffix as *-(z)zjoni* rather than *-zzjoni*.

There are a number of cases where *-Vr* and *-(z)zjoni* forms share the same base. Examples include, *istallar* and *istallazzjoni* 'installation', both of which are related to *i(n)stalla* 'to install'. The difference in meaning is not always clear, though generally it appears that the *-Vr* version refers to a process or event

(close to English *-ing* formation, as in 'installing'), while *-(z)zjoni* can refer to either a process/event or an entity (similar to English *installation*, i.e., the result of an installation process; see Ellul 2016 and references therein for a resultative analysis of such forms). This observation is not without exceptions, however, as shown by examples such as *armar* 'decoration/decorating' and *tellar* 'panel beater', neither of which have a corresponding *-(z)zjoni* form. In any case, though there are cases where both an *-Vr* and a *-(z)zjoni* form coexist with the same base, most are found exclusively in either one or the other form. Table 3 gives examples of bases which nominalise exclusively in one or the other form.

Table 3: Bases which nominalise using *-Vr* or *-(z)zjoni* , but not both.

-Vr Nominalisation	*-(z)zjoni* Nominalisation	Gloss
ibbukkjar	*ibbukkjazzjoni	booking
depożitar	*depożitazzjoni	depositing
ittestjar	*testazzjoni	testing
*traduttar	traduzzjoni	translation
*assumar	assunzjoni	assumption
*affaxxinar	affaxxinazzjoni	fascination

4.2 Direct or indirect borrowing?

In the previous section, we observed that certain affixes borrowed from Italian may be cases of what Seifart (2015) calls indirect borrowing, since they are used on native stems. Here, we revisit this question in connection with *-Vr* and *-(z)zjoni* .

There are various cases of *-Vr* being used on stems of Arabic origin, as shown below.

(4) (Suffix *-Vr* applied to stems of Arabic origin; personal knowledge)
ittama ↔ *ittamar* *tkaża* ↔ *tkażar*
to hope ↔ hope be shocked ↔ shock

By contrast, the suffix *-(z)zjoni* does not seem to be used with stems of Arabic origin. In our corpus data (§5), we have been unable to identify a single case, nor does our intuition as native speakers suggest any examples. However, there are several cases where the affix is used with stems of non-Romance origin, especially English, as shown below.

(5) (Suffix *-(z)zjoni* applied to stems of English origin; Michael Spagnol, pc.)
esplojta ↔ *esplojtazzjoni* *immoniterja* ↔ *moniterizzazzjoni*
to exploit ↔ exploitation to monitor ↔ monitoring

More clearly 'local' in origin are formations where *-(z)zjoni* is applied to lexemes ending in *-izza* (roughly, the equivalent of English *-ise* or Italian *-izzare*), which are in a derivational relationship to a proper name. These complex formations are frequently candidates for nominalisation using *-(z)zjoni* .[1]

(6) (Suffix *-(z)zjoni* applied to proper names; personal knowledge)
Xarabank ↔ *Xarabankizzazzjoni* *Dubaj* ↔ *Dubajizzazzjoni*
Xarabank ↔ Xarabankisation Dubai ↔ Dubaification

It is possible that rather than being clear-cut cases of direct or indirect borrowing, the suffixes under consideration should more accurately be placed somewhere along the continuum between these two extremes. This can be done by weighing the empirical evidence for indirect borrowing in the two cases, using the following criteria provided by Seifart (2015: p. 513):

1. A set of complex loanwords with the borrowed affix share a meaning component;
2. There exists a set of pairs of loanwords, with one element of each pair with the affix and one without, with constant, recognisable changes in meaning between them;
3. Within pairs of complex loanwords and their corresponding simplex loanwords, the former have a lower token frequency.

Of these, the first criterion seems easily satisfied by both *-Vr* and *-(z)zjoni* , insofar as the many forms with these borrowed affixes do share a meaning component, as well as a formal relationship by virtue of having the same nominalising suffix. It is the second and third criteria that are clearly testable. Below, we present a quantitative analysis of the productivity of these affixes, and then turn to the evidence for or against these two criteria. As noted in §1, we view the corpus-based investigation of in/direct borrowing and its implications for the parseability of forms (Hay & Baayen 2001) as complementary to the question of productivity.

[1] In the example below, Xarabank is the name of a discussion programme on Maltese national television.

5 An empirical investigation

We now turn to a quantitative analysis of the productivity of the two nominalisation affixes under discussion. We take a corpus-based approach to address the following question: *How productive are -Vr and -(z)zjoni nominalisations in Maltese, that is, to what extent are the two processes likely to contribute novel forms?* We then turn to the criteria for indirect borrowing, and the extent to which we find evidence for them in the two cases.

In quantifying productivity, we take inspiration from the statistical account offered by Baayen (Baayen 1994; 2009) and developed in subsequent work (for example, Lüdeling et al. 2000 and Pustylnikov & Schneider-Wiejowski 2010). We first discusss Baayen's theoretical framework, before describing the data used for this analysis.

5.1 Baayen's productivity measures

Baayen (2009) distinguishes between three conceptions of productivity. The *realised productivity* (RP) of a morphological process is defined as the number of types in a corpus that have been formed using this process. By contrast, *expanding productivity* (P^*) refers to the extent to which a process contributes to the growth rate in the total vocabulary, as reflected in a particular corpus. It is computed as the proportion of hapax legomena formed via the process in question, out of the total number of hapaxes in the corpus. As such, it is intended to reflect the number of 'novel' forms that the process has contributed, where 'novel' is operationally defined as a one-off occurrence, under the assumption that a word with a frequency of 1 is potentially a newly coined form.

Both RP and P^* are strongly dependent on corpus size, since both the total vocabulary size and the number of hapax legomena tend to grow – albeit asymptotically – with corpus size (see Baroni 2008 for discussion). Baayen's final measure of productivity – referred to as *potential productivity* or *category-conditioned productivity* and denoted P – focusses instead on the proportion of hapax legomena formed using the process in question, out of the total number of tokens that are formed using that process. This is less susceptible to variation due to corpus size, since it is related to the total number of tokens arising from a given process. P is usually taken to be the most reliable quantitative indicator of productivity out of the three. It is also interpreted as an indicator of the rate at which the morphological process could be used to create novel or 'potential' forms. In particular, the number of one-off occurrences out of the total number of tokens formed using a process should give us some indication of the relative prevalence of coinages or new usages.

There have been some criticisms of the use of P as formulated by Baayen (2009). In particular, Gaeta & Ricca (2006) argue that because P relies on the number of tokens created via a morphological process, it tends to underestimate the productivity of processes with high token frequency, while overestimating the productivity of forms with lower token frequency. For example, Gaeta & Ricca (2006) find that the Italian nominalising suffix *-tore* would be estimated as much more productive than its feminine counterpart *-trice*, which has a lower token frequency. As a corrective measure, Gaeta & Ricca (2006) suggest using the *variable-corpus* approach, in which morphological processes are compared for their productivity at varying token frequencies. By this argument, given two processes A and B, with token frequencies N_A and N_B such that $N_A < N_B$, the comparison of P would be more meaningful if N_A is used in the denominator. This method has also been used by Saade (2016) for his comparison between Maltese and Italian derivations.

While these arguments are well-taken, they are nevertheless subject to counter-arguments. In particular, since P is by definition estimated relative to token frequency, it is to be expected that as a process becomes more frequent and exhausts its domain of potential application, its productivity will be reduced. A similar argument has been put forward by Baayen (2009).

In the present paper, we will stick to the original proposals made by Baayen for the estimation of P, which we take to be indicative of the likelihood that a morphological process will yield novel forms in future. However, we also take the following additional methodological steps:

1. We estimate productivity over multiple, equal-sized corpus samples. This does not imply that we restrict the denominator in the estimation of P to the minimum token frequency for *-Vr* and *-(z)zjoni* ; rather, we obtain multiple measures that also allow the investigation of the effect of increasing corpus size.

2. We consider both vocabulary growth and productivity for *-Vr* and *-(z)zjoni* as a function of increasing corpus size, as well as over the entire corpus.

3. We consider the correlation between the three measures of productivity.

Before turning to the analysis, we give a description of the data used.

5.2 Corpus data

The present analysis draws on data from the Korpus Malti v2.0 Beta, a corpus of ca. 125 million tokens developed and distributed as part of the Maltese Language

Table 4: Distribution of texts in the MLRS Korpus Malti v2.0 Beta, after Gatt & Čéplö (2013).

Text type	Number of tokens
Journalistic texts	68.800.000
Parliamentary debates	43.400.000
Belles lettres	375.000
Academic texts	170.000
Legal texts	4.800.000
Religious texts	403.700
Speeches	18.000
Web pages (blogs, Wikipedia articles, etc)	6.500.000
Miscellaneous other texts	123.000

Resource Server (MLRS).[2] The corpus is tagged with part of speech information, and contains texts from a variety of genres, as shown in Table 4 (Gatt & Čéplö 2013).

For the purposes of our analysis, we took 15 random samples of 1000 sentences each from the corpus. The decision to use multiple samples rather than conduct a single analysis on the corpus as a whole was motivated by three factors. First, using relatively small samples facilitates the manual pruning of false positives from search results (a well-known problem in analyses of morphological productivity; see Pustylnikov & Schneider-Wiejowski 2010). In the present case, for example, false positives include lexemes which end in *-ar* but are not derived nominals, such as *mar* 'go'; *parpar* 'scarper'; and *għargħar* 'deluge'.[3] Second, the ability to compute the productivity measures over multiple samples provides us with multiple data points, enabling a correlational analysis between the productivity measures, as presented in §5.5 below. Finally, multiple samples also allow the estimation of vocabulary and productivity curves over samples of increasing size, as presented in §5.4 and §5.5 below.

[2] http://mlrs.research.um.edu.mt

[3] There seems to be no straightforward way of automating the detection of false positives based on simple criteria such as length. While it would be possible to train a classifier to distinguish true from false positives, it was deemed better, on balance, to apply manual filtering, since the accuracy of automatic classification would in any case probably not reach 100%, and furthermore, an investigation of the features necessary to distinguish true and false positives is well beyond the scope of the present paper.

Each of the 15 random samples was pre-processed as follows:

1. Extraction of tokens tagged as nouns and ending in *-Vr* or *-(z)zjoni*: for the former, we restricted attention to the form *-ār* since the alternative form seems to be restricted to only a few types (cf. the discussion in §4 and the work of Camilleri 1993);

2. Manual pruning of false positives, specifically, nouns with these endings that are not the outcomes of the derivational processes under discussion (e.g. *għar* 'cave', which is not a derived nominal);

3. Extraction, from each sample, of the frequency distribution of types belonging to each process.

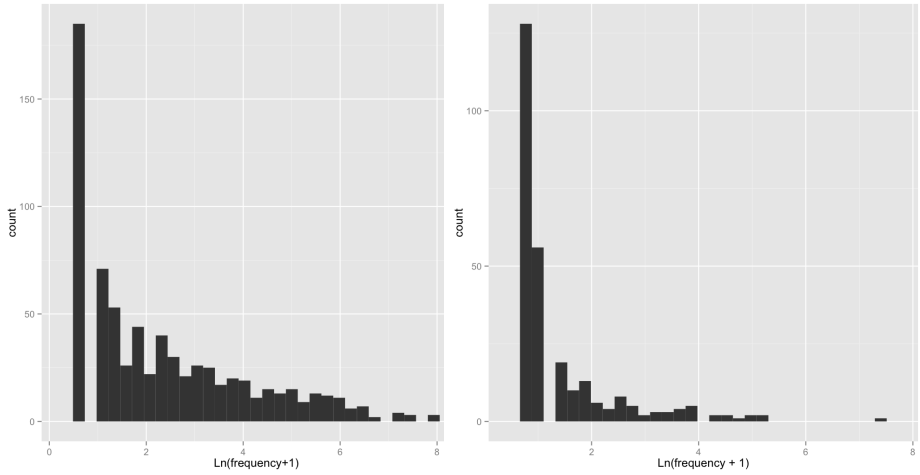
5.3 The distribution of *-Vr* and *-(z)zjoni* nominalisations

Table 5 gives an overview of the main characteristics of the samples under analysis, as well as the mean size and vocabulary, number of hapax legomena, and frequencies of *-ar* and *-zjoni* derivations overall.

A few observations are worth making at the outset. First, the 15 corpus samples are relatively homogeneous, with sizes ranging from 257,586 to 264,482 tokens and vocabulary sizes ranging from 23,788 to 24,345. Second, it is immediately clear that the incidence of *-(z)zjoni* nominalisations is far higher than that of *-Vr*

Table 5: Basic statistics for the samples used in the analysis. All figures average over the 15 random samples of 1000 sentences each.

	Mean	St. Dev	Min	Max	Median
Tokens	260,533	1655	257,586	264,482	260,186
Types	24,092	169	23,788	24,345	24,132
Hapaxes	12,611	155	12,340	12,872	12,617
Tokens: *-zjoni*	3,519	137	3,234	3,712	3,512
Types: *-zjoni*	325	20	305	382	325
Hapaxes: *-zjoni*	114	17	93	161	109
Tokens: *-ar*	256	21	227	288	258
Types: *-ar*	61	6	49	73	62
Hapaxes: *-ar*	35	5	25	43	35



(a) *-(z)zjoni* nominalisations (b) *-Vr* nominalisations

Figure 1: Frequency histograms for *-Vr* and *-(z)zjoni* nominalisations. Frequencies are plotted on a logarithmic scale on the x-axis, adding 1 to avoid zero frequencies for hapax legomena.

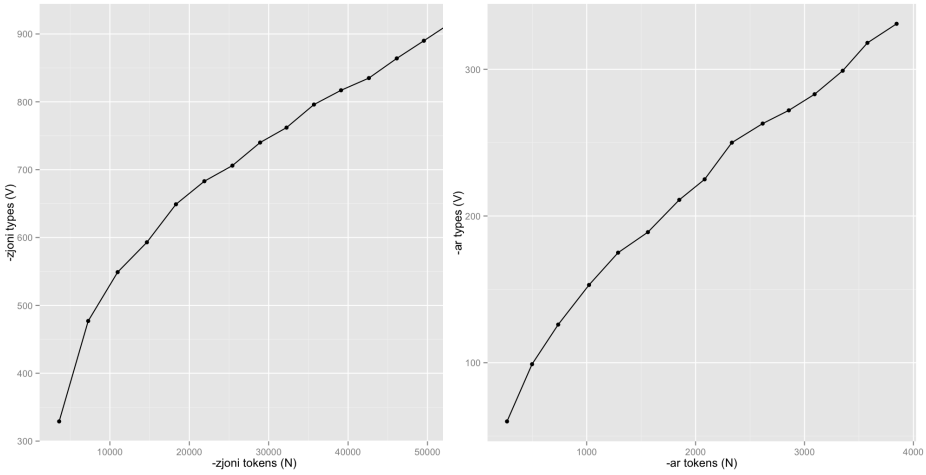
nominalisations: On average, there are 13 times as many tokens of the former as there are of the latter, and 5 times as many types.

Figure 1 displays the type frequency histograms, on a logarithmic scale, for the two processes. Interestingly, *-Vr* nominalisations tend to exhibit a much steeper drop in frequency, and a more uneven distribution, with a substantial gap between the hapax legomena and the next highest frequency. By contrast, *-(z)zjoni* nominals tail off more evenly. In general, not only are there more *-(z)zjoni* types, but there are more types within each frequency interval.

5.4 Vocabulary growth

A useful way to obtain a preliminary indication of the productivity of the nominalisation processes *-Vr* and *-(z)zjoni* is to look at their vocabulary growth curves. These display the size of the vocabulary (that is, the number of types V) as a function of increasing numbers of tokens (denoted N), generated using those processes.

As Lüdeling et al. (2000) note, a relatively unproductive process will tend to exhibit a shallow or asymptotic $N \times V$ curve, with vocabulary size no longer increasing as tokens increase in number. This means that beyond a certain point, as tokens increase, there tend not to be so many instances of novel, previously



(a) *-(z)zjoni* nominalisations (b) *-Vr* nominalisations

Figure 2: Vocabulary growth curves for *-Vr* and *-(z)zjoni* as a function of increasing number of tokens. Types are plotted on the y-axis, with tokens on the x-axis.

unattested types. By contrast, the more productive a process is, the steeper the $V \times N$ curve is expected to be.

Vocabulary growth curves were obtained for both *-Vr* and *-(z)zjoni* nominalisations by computing the number of different types over increasingly large samples, obtained by cumulatively merging the data from our 15 random samples and recomputing the token and vocabulary counts at each step. The curves are displayed in Figure 2. The vocabulary growth curves have a similar shape, showing a steep increase in both cases. This provides some *prima facie* evidence that both processes are productive, despite the much higher relative frequency of *-(z)zjoni* formations compared to *-Vr* formations noted in §5.3 above. As the histograms in Figure 1 confirm, this is due to the greater number of high-frequency types in the case of *-(z)zjoni* , also shown by the more even shape of the distribution in Figure 1a. The evidence therefore suggests that the productivity of these processes is independent of their absolute frequency.

5.5 Productivity analysis

We turn now to the quantification of productivity of the two derivational processes, using the measures proposed by Baayen (2009). For the purposes of this part of the analysis, the three measures, RP, P^* and P, were computed sepa-

Table 6: Productivity measures for the two derivational processes. All figures average over the 15 samples; numbers in parentheses are standard deviations.

	Proportional RP	P*	P
-*Vr* nominalisations	0.00245 (0.0002)	0.00275 (0.0004)	0.136 (0.02)
-*(z)zjoni* nominalisations	0.0135 (0.0008)	0.00904 (0.001)	0.033 (0.005)

rately for each sample. This gives us 15 data points, which can be used to correlate the three measures. Table 6 summarises the findings, showing the mean of each of the three measures, across samples. In these figures, RP is estimated as the proportion of types out of all the types in the sample.

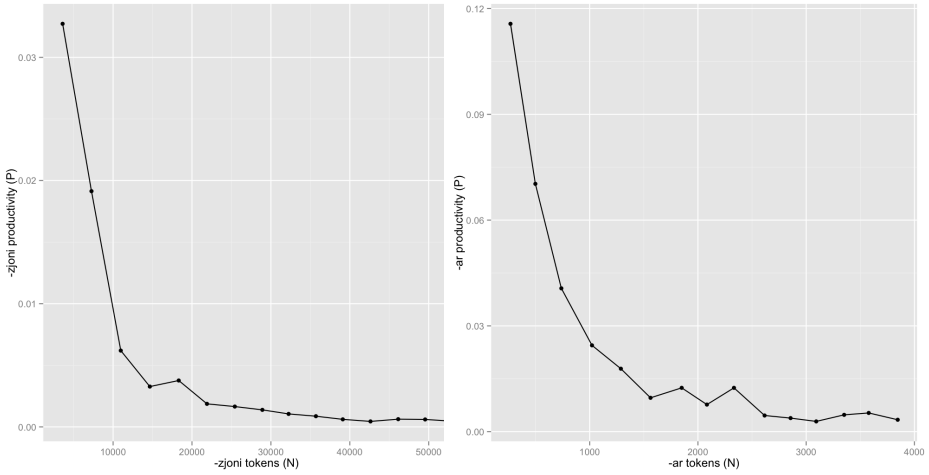
Although, as noted in the introduction to this subsection, the three productivity measures are intended to reflect different perspectives on productivity, we nevertheless expect them to be correlated since they each depend on the overall vocabulary size (or on that part of the vocabulary that consists of one-off occurrences, or hapax legomena).

The three productivity measures are highly positively correlated, as shown in Table 7. One partial exception is the correlation between RP and P for *-Vr* nominalisations, which is only marginally significant at $p \approx 0.06$. Over all, however, there is systematic covariation between the three quantitative perspectives on productivity.

However, what is perhaps most interesting from the perspective of this analysis is that while *-(z)zjoni* exhibits greater realised productivity (RP) and expanding productivity (P^*) than *-Vr* does, the trend is reversed where potential

Table 7: Pearson's correlation coefficients between the productivity measures for each nominalisation process.
* indicates that the correlation is significantly different from 0 a $p \leq 0.001$; † indicates that the correlation approaches significance at $p \approx 0.06$.

	-Vr Nominalisations		*-(z)zjoni* Nominalisations	
	P*	**P**	**P***	**P**
Proportional RP	0.80*	0.50†	0.94*	0.87*
P*	–	0.86*	–	0.97*



(a) *-(z)zjoni* nominalisations (b) *-Vr* nominalisations

Figure 3: Productivity (P) of *-Vr* and *-(z)zjoni* as a function of increasing numbers of tokens.

productivity (P) is concerned, as shown in Table 6 above: Potential productivity is greater for *-Vr* than for *-(z)zjoni* . On the basis of this data, then, *-Vr* would be expected to contribute a greater proportion of new vocabulary than *-(z)zjoni* . This is interesting in light of the fact – evident from Figure 2b – that unlike in the case of *-(z)zjoni* , the distribution of *-Vr* types shows a gap between hapaxes and the preceding frequency intervals in the histogram. Taken together, the evidence points towards *-Vr* having a tendency to be used to create novel types, which are reflected as 'one-offs' in the corpus.

As with the vocabulary growth curves in Figure 2 above, we successively merged samples to create larger corpora and re-estimated the number of hapaxes for *-Vr* and *-(z)zjoni* , estimating P as the total number of tokens formed via a particular process increases. The resulting curves are displayed in Figure 3. As expected, both processes show a decrease in P with increasing number of tokens. This is expected, since the proportion of hapaxes tends to decrease as corpus size grows. However, the potential productivity of *-(z)zjoni* drops to a value close to 0 more steeply than does that of *-Vr* .

5.6 Evidence for indirect borrowing

We now turn to the two (out of three) criteria for indirect borrowing outlined by Seifart (2015) and singled out in §4.2. Recall, from our discussion in that section,

that our concern is to determine whether *on balance* the evidence points towards these affixes being indirectly borrowed.

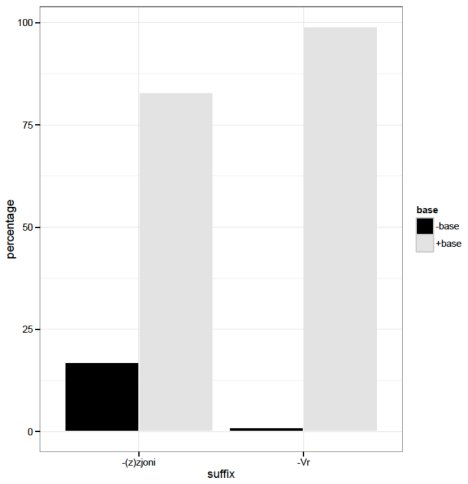
This part of the analysis proceeded as follows:

1. We determined, for each lexeme in our sample, its corresponding verbal baseform, if any. For example, the nominalisation *sparar* 'shooting' has a corresponding verbal baseform *spara* 'shoot'. Similarly, *informa* 'inform' corresponds to *informazzjoni*. On the other hand, a number of nominalisations do not have corresponding baseforms in Maltese. For example, there is no verb derivationally related to *devozzjoni* 'devotion'; *demozzjoni* 'demotion'; or *inġunzjoni* 'injunction', though these nominals are all attested in the corpus.

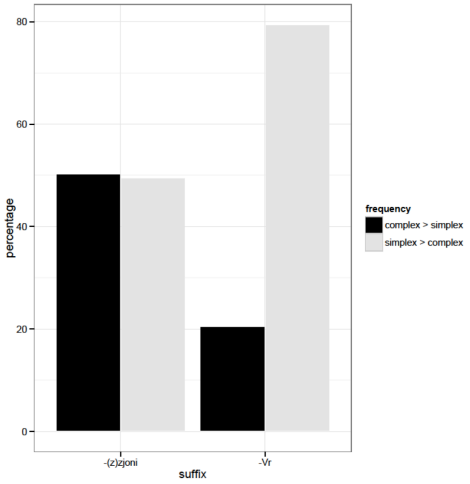
2. We compared the number of types formed with *-(z)zjoni* and *-Vr*, across the entire corpus (i.e. combining all 15 samples), which have corresponding simplex (verb) forms. This sheds light on the evidence for Seifart's second criterion, which stipulates that in case of indirect borrowing, loanwords will typically occur in pairs, where one element has the affix and one does not. The results are displayed in Figure 4a.

3. We also compared the token frequency of forms with and without the affix (i.e. complex and simplex verb forms). Here, we are interested in the number of types formed with a given affix which have lower token frequency than their corresponding simplex forms, as predicted by Seifart's third criterion. For this part of the analysis, we therefore only focus on that subset of the nominalisations identified in the previous step for which corresponding simplex forms are attested. We used the whole of Korpus Malti v2.0. Using the frequency list for this corpus, we extracted the frequency of the nominalisations and that of their corresponding verb forms. Given that verbs in Maltese can be inflected for person, number and gender, and that, furthermore, they can take a set of enclitic object pronouns, the verb forms were identified heuristically by finding all the lexemes in the frequency list which contained the verb stem as a substring, excluding the *-Vr* or *-(z)zjoni* nominalisations themselves.[4] The results are displayed in Figure 4b.

[4]This heuristic therefore only gives an approximate estimate of the verb frequency. False negatives are possible for those words which are misspelled in the corpus, as when an author uses *iccekkja* instead of *iċċekkja* 'check'. False positives are in principle possible insofar as a word may have the verb stem as a substring, but be unrelated to it. Though possible, this is relatively unlikely, given that the verb forms have a fairly clear structure and are regular, with little stem allomorphy.



(a) Proportions of *-Vr* and *-(z)zjoni* nominalisations that have corresponding verbal baseforms. Dark bars represent types with no corresponding baseform; light bars reflect types with a corresponding baseform.



(b) Comparison of the frequency of complex and simplex forms. Dark bars: complex > simplex; light bars: simplex > complex.

Figure 4: Evidence for indirect borrowing: Nominalisations and corresponding baseforms

Two observations can be made from this analysis. First, as far as Seifart's second criterion goes, both nominalisation processes evince a majority of types with corresponding simplex forms. However, this is far more likely with *-Vr* nominalisations (ca. 99% of cases) than *-(z)zjoni* nominalisations (ca. 83% of cases). This suggests that there are more cases of *-(z)zjoni* lexemes which were borrowed wholesale, rather than produced 'online' from stems by native speakers. This conclusion is strengthened by the apparent absence of *-(z)zjoni* forms involving native Semitic stems, observed in §4.2.

Second, as far as token frequency is concerned, *-(z)zjoni* hardly satisfies Seifart's third criterion: with this form, the proportion of cases where the simplex form is more frequent than the nominalised form is roughly equal to the proportion of cases where the opposite holds (both are around 50%). By contrast, over 79% of types formed using *-Vr* are less frequent than their simplex forms.

On balance, therefore, the evidence for indirect borrowing is much more clear in the case of *-Vr* than *-(z)zjoni* .

5.7 Summary

This corpus-based analysis sheds light on the productivity of the two nominalisation processes from two different perspectives. First, the productivity analysis suggests that both *-Vr* and *-(z)zjoni* are productive to some degree. This is reflected both by their vocabulary growth curves and by their non-zero estimates for potential productivity (P). At the same time, it is noteworthy that the process whose formations are most frequently attested – namely, *-(z)zjoni* – turns out to have a lower potential productivity, despite its apparently higher realised (RP) and expanding productivity (P^*). As noted above, the latter two measures are more strongly dependent on corpus size (Baayen 2009).

What could account for the higher P measure for *-Vr* , when *-(z)zjoni* has higher RP and P^*? One possible reason, alluded to in §2, is that, despite the larger number of attested *-(z)zjoni* types, there are also more forms without corresponding simplex forms, because a larger proportion of these types was imported wholesale, so that these types are not derivationally related to an attested verb. Hence, the more corpus-dependent (as opposed to category-conditioned) productivity measures would be inflated by a greater proportion of types that are in fact not derivationally related to bases in the native speaker's mental lexicon.

The second part of the analysis, focussing on the criteria outlined by Seifart (2015) for indirect borrowing, strengthens this position. Specifically, we find that *-Vr* lexemes are more likely to have corresponding simplex forms. Furthermore, a

comparison of the frequency of complex and simplex forms shows that the latter are more likely to be used with greater frequency in the *-Vr* case, compared to *-(z)zjoni* . This provides further evidence for wholesale importation of *-(z)zjoni* forms, suggesting that the greater productivity of *-Vr* is in part due to its large-scale re-use on novel, including native, stems, possibly following a process of reanalysis of forms originally imported from Italian, after which the affix became available for use on a broader domain.

In any case, to the extent that the domains of application of the two derivational processes overlap (cf. §4 above), the figures for potential productivity (P) suggest that there will be a greater preference among speakers for forming nominalisations using *-Vr* rather than *-(z)zjoni* in the future. Clearly, this conclusion can only be tentatively reached on the basis of corpus data, especially since such data, by definition, is 'historical' and restricted to already-attested, rather than potential forms.

6 Conclusions

This paper began with an outline of morphological derivation in Maltese, couched within a theoretical framework that is agnostic as to the procedural nature of the derivational process, focussing instead on the relation between two lexemes. Following an outline of both Semitic and Romance derivational processes, we focussed on two derivational suffixes – *-Vr* and *-(z)zjoni* – which appear to share a number of semantic and distributional characteristics. A corpus-based analysis showed that one of them, namely *-Vr* , is likely to emerge as more productive in the long-term. The evidence further points to a greater likelihood that *-Vr* was indirectly borrowed into Maltese, coming to be used on a broader range of stems, including native stems.

The snapshot provided by the present analysis opens up various avenues for future research. An important one is the in-depth analysis of a greater variety of derivational processes, with a view to providing a deeper understanding of derivational morphology in contemporary Maltese as well as gaining a better understanding of the extent to which the domains of such processes overlap. A second important direction for future work is the exploitation of different methodological tools. As the present paper showed, corpus analysis can provide substantial insights into questions related to morphological productivity. However, we believe that such analyses need to be complemented by experimental techniques, which can shed a more direct light on the processing implications of the trends observed in corpora.

Acknowledgements

This work forms part of the project *Morphological Productivity and Language Change in Maltese: Corpus-based and experimental evidence*, supported by a grant from the University of Malta Research Fund. We thank Michael Spagnol, Manwel Mifsud, Benjamin Saade and two anonymous reviewers for helpful comments on earlier drafts of this paper.

References

Aronoff, Mark & Frank Anshen. 1998. Morphology and the lexicon: Lexicalization and productivity. In Andrew Spencer & Arnold Zwicky (eds.), *The handbook of morphology*, 237–247. Oxford: Blackwell.

Aronoff, Mark & Kirsten Fudeman. 2011. *What is Morphology?* 2nd edn. Oxford: Wiley-Blackwell.

Aronoff, Mark & Roger Schvaneveldt. 1978. Testing morphological productivity. *Annals of the New York Academy of Sciences* 318. 106–114.

Baayen, R. Harald. 1994. Productivity in language production. *Language and Cognitive Processes* 9(3). 447–459.

Baayen, R. Harald. 2009. Corpus linguistics in morphology: Morphological productivity. In Anke Lüdeling & Merja Kytö (eds.), *Corpus linguistics. An international handbook*. Chap. 43, 900–919. Berlin: de Gruyter.

Baroni, Marco. 2008. Distributions in text. In Anke Lüdeling & Merja Kytö (eds.), *Corpus linguistics: An international handbook*, chap. 37, 803–821. Berlin: Mouton de Gruyter.

Bauer, Laurie. 2001. *Morphological Productivity*. Cambridge: Cambridge University Press.

Borg, Albert & Marie Azzopardi-Alexander. 1997. *Maltese*. London: Routledge.

Brincat, Joseph. 2011. *Maltese and other languages: A linguistic history of Malta*. Malta: Midsea Books.

Brincat, Joseph & Manwel Mifsud. 2016. Maltese. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *Word-formation: An international handbook of the languages of Europe (vol. 5)*, 3349–3366. Berlin: De Gruyter Mouton.

Bybee, Joan. 1995. Regular morphology and the lexicon. *Language and Cognitive Processes* 10(5). 425–455.

Camilleri, Edel. 1993. *Is-Suffissi Nominali Rumanzi fil-Malti (Romance nominal suffixes in Maltese)*. University of Malta B.A. Honours Dissertation.

Comrie, Bernard & Michael Spagnol. 2016. Maltese loanword typology. In Gilbert Puech & Benjamin Saade (eds.), *Shifts and patterns in maltese*, 315–330. Berlin: De Gruyter Mouton.

Cutler, Anne. 1980. Productivity in word formation. In *Papers from the sixteenth regional meeting, Chicago Linguistic Society*, 45–51. Chicago, Ill: CLS.

Dressler, Wolfgang. 2003. Degrees of grammatical productivity in inflectional morphology. *Rivista di Linguistica* 15(1). 31–62.

Drewes, A.J. 1994. Borrowing in Maltese. In Peter Bakker & Maarten Mous (eds.), *Mixed languages: 15 case studies in language intertwining*, 83–111. Amsterdam: Ifott.

Ellul, Leanne. 2016. *In-Nomi Verbali fil-Malti (Verbal nouns in Maltese)*. University of Malta Masters Thesis.

Fabri, Ray. 2010. Maltese. *Revue Belge de Philologie et d'Histoire* 88(3). 791–816.

Fabri, Ray, Michael Gasser, Nizar Habash, George Kiraz & Shuly Wintner. 2013. Linguistic Introduction. In Zitouni, I. (ed.), *Semitic language processing*, 16–21. Berlin: Springer.

Gaeta, Livio & Davide Ricca. 2006. Productivity in Italian word formation: A variable-corpus approach. *Linguistics* 44(1). 57–89. DOI:10.1515/LING.2006.003

Gatt, Albert & Slavomír Čéplö. 2013. Digital corpora and other electronic resources for Maltese. In *Proceedings of the International Conference on Corpus Linguistics*. Lancaster, UK: University of Lancaster.

Haspelmath, Martin & Andrea Sims. 2010. *Understanding Morphology*. 2nd edn. London: Routledge.

Hay, Jennifer & R. Harald Baayen. 2001. Parsing and productivity. In Geert Booij & Jaap van Marle (eds.), *Yearbook of morphology 2001*, 203–235. Dordrecht: Kluwer. DOI:10.1007/978-0-387-93837-0

Hoberman, Robert D. & Mark Aronoff. 2003. The verbal morphology of Maltese: From Semitic to Romance. In J. Shimron (ed.), *Language processing and acquisition in languages of Semitic, root-based morphology*, chap. 3, 61–78. Amsterdam: John Benjamins Publishing Company.

Lüdeling, Anke, Stefan Evert & Ulrich Heid. 2000. On Measuring Morphological Productivity. In Werner Zühlke & Ernst G. Schukat-Talamazzini (eds.), *KONVENs-2000 sprachkommunikation*, 57–61. Germany: VDE-Verlag.

Mifsud, Manwel. 1995a. *Loan verbs in Maltese: A descriptive and comparative study*. Leiden: Brill.

Mifsud, Manwel. 1995b. The productivity of African in Maltese. In Joseph Cremona, Clive Holes & Geoffrey Khan (eds.), *Proceedings of the 2nd International*

Conference of the International Association for African Dialectology (AIDA), 151–160. Cambridge, UK: University Publications Centre.

Plag, Ingo. 2004. Productivity. In Bas Aarts & April McMahon (eds.), *Handbook of English linguistics*, 537–556. Oxford: Blackwell.

Pustylnikov, Olga & Karina Schneider-Wiejowski. 2010. Measuring morphological productivity. *Studies in Quantitative Linguistics* 5. 1–9.

Saade, Benjamin. 2016. Assessing productivity in contact: Italian derivation in Maltese. *Linguistics* to appear.

Seifart, Frank. 2015. Direct and indirect affix borrowing. *Language* 91(3). 511–532. DOI:10.1353/lan.2015.0044

Spagnol, Michael. 2013. *A Tale of Two Morphologies: Verb Structure and Argument Alternations in Maltese*. University of Konstanz PhD Thesis.

Spencer, Andrew. 1991. *Morphological theory*. Oxford: Blackwell.

Spencer, Andrew & Arnold Zwicky. 1998. *The Handbook of Morphology*. Oxford: Blackwell.

Twist, Alina. 2006. *A Psycholinguistic Investigation of the Verbal Morphology of Maltese*. University of Arizona, Tucson (AZ) PhD Thesis.

Ussishkin, Adam, Colin Dawson, Andrew Wedel & Kevin Schluter. 2015. Auditory masked priming in Maltese spoken word recognition. *Language, Cognition and Neuroscience* January 2015. 1–20.