

# Probing the Nature of Roots through Language Contact

Artemis Alexiadou and Terje Lohndal

## 1. Introduction

The nature of roots and their role in linguistic theory have long preoccupied linguists of various theoretical orientations. A number of decompositional theories have proposed that roots are among the basic building blocks of word formation. For instance, within Distributed Morphology (e.g., Halle & Marantz 1993, Marantz 1997, Embick 2015), a root combines with a categorizer to create a category. If the categorizer is verbal, a verb is created, whereas a noun is created if it is nominal. This is illustrated in (1).



On this view, lexical categories are not primitives, they are the result of a syntactic derivation. Rather, roots and functional categorizers are primitives.

Despite the prevalence of roots, theories do not agree on their specific characteristics (see, e.g., the collection of essays in Alexiadou, Borer & Schäfer 2014). In the present paper, we will approach the nature of roots from a different angle, namely through the lens of language contact. Based on two case studies of language mixing word-internally and in compounds, we claim that studying word formation processes where properties of more than one language are involved offers a unique window onto the basic word formation mechanisms and the nature of roots. Specifically, we argue that data from language contact support the following view of roots: i) roots are devoid of categorial information (cf. Marantz 1997, Embick 2015), ii) roots merge with (overt or covert) categorizers in the syntax (Marantz 1997, Embick 2015, Alexiadou & Lohndal 2017; though see Borer 2014), iii) more generally, the realization of roots and their functional vocabulary as identified in work on mixing informs theories of what a word is, how it is built, and how concepts are lexicalized across languages.

This paper is organized as follows. Section 2 provides some background. Sections 3 and 4 provide the two case studies of word-internal language mixing and language mixing in compounds, respectively. Section 5 provides a synthesis of the two case studies. Concluding remarks are made in section 6.

## 2. Background

Language mixing involves lexical items and grammatical features from two (or more) languages that appear in one sentence (Muysken 2000). Pioneering work on the typology of language mixing has been done by Muysken (2000, 2013), which largely focuses on word-level units and beyond. Relatively little work has been done on the mechanisms involved in language mixing below the level of the word, mostly because it was argued to only occur in limited environments. For instance, Poplack (1980) and especially

---

\* We are grateful to the audience at WCCFL for their helpful questions and comments. This paper is part of the international research project *MultiGender* at the Centre for Advanced Study at the Norwegian Academy of Science and Letters in Oslo. AL 554/8-1 (Alexiadou) is also hereby acknowledged.

Artemis Alexiadou, Humboldt University of Berlin & ZAS, artemis.alexiadou@hu-berlin.de. Terje Lohndal, NTNU – Norwegian University of Science and Technology & UiT The Arctic University of Norway, terje.lohndal@ntnu.no.

Poplack & Sankoff (1981) argued that such mixing is conditional upon phonological integration: Mixing can only take place if a morphological ending is integrated into the language of the lexical form. Their constraint allows forms such as in (2), but not those in (3).

- |        |   |    |  |
|--------|---|----|--|
| (2) a. | flip-eando<br>flip-ing<br>'flipping' (Sankoff & Poplack 1981: 5)) | b. | parqu-eando<br>park-ing<br>'parking' (MacSwan 2005: 7)         |
| (3) a. | *eat-iendo<br>eat-ing<br>'eating' (Poplack 1980: 586)             | b. | *run-eando<br>run-ing<br>'running' (Sankoff & Poplack 1981: 5) |

In (2), *flip* and *parqu* have been adapted to Spanish phonology, whereas this is not the case for *eat* and *run* in (3). Since the seminal work done in the early 1980's, scholars have often distinguished between (2) and (3) in terms of borrowing and code-switching. If a word is phonologically integrated, we are dealing with borrowing, and if there is no phonological integration, it is an instance of code-switching. However, this distinction is controversial, and it has been argued that from a formal point of view, distinguishing between borrowing and code-switching is futile (cf. Alexiadou & Lohndal 2018 and the review of the literature there). Here, we will use the term 'language mixing' which is neutral as to the distinction between borrowing and code-switching.

In recent years, the role of language mixing inside words has gained prominence, as witnessed in publications such as González-Vilbazo & López (2011), Callies & Stolz (2016), Áfarli (2015), Alexiadou & Lohndal (2018), Grimstad et al. (2018), Riksem et al. (2019), Alexiadou (2020), and López (2020) (see also Wohlgemuth 2009 from a different perspective). These publications argue that mixing between two or more languages can be used to identify the basic units that are involved in both monolingual and bilingual word formation. In particular, it is argued that data from language mixing support a decompositional view of morphology, whereby morphemes are the realizations of abstract syntactic features. Furthermore, this work has argued that these morphemes are assembled syntactically, which is to say that words are composed by way of the same principles as sentences.

Even though previous research has argued for a decompositional approach to morphology, it has rarely explicitly addressed the nature of roots. Instead, roots have often been adopted without a detailed assessment of their properties. The current contribution seeks to fill this lacuna by demonstrating what data from language contact can tell us about the nature of roots. We will focus on two case studies, namely word-internal language mixing, and language mixing in compounds.

### 3. Case Study 1: Word-internal language mixing

In the first case study, we will consider language mixing internal to words. A useful illustration of how word-internal language mixing serves to identify the basic building blocks of words can be seen in (4). (4a) illustrates word-internal verbal mixing in Greek-German whereas (4b) is from Cypriot-Greek-English (see also Alexiadou 2017).

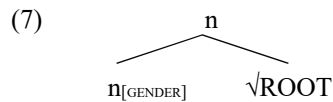
- |        |  |    |   |
|--------|--|----|---|
| (4) a. | <b>skan</b> -ar-o<br>scan-AFF-1SG<br>'I am scanning.' (Alexiadou 2011) | b. | <b>kansel</b> -ar-o<br>cancel-AFF-1SG<br>'I am cancelling' (Gardner-Chloros 2009: 50-1) |
|--------|--|----|---|

Here a Greek affix attaches to the German and/or English root. A dedicated affix, *-ar-*, is used to verbalize the root. This is the default verbalizer in cases of mixing. We can model this as an uncategorized non-native root that is merged with a Greek verbal categorizer, making the resulting unit a verb. If roots are not recognized, it is not equally straightforward how to characterize this pattern. Note that it is always a German/English root that combines with a Greek affix, and speakers reject the combination of a Greek root with a German inflection. We won't discuss the latter asymmetry here; see Alexiadou (2017) for an analysis.

Mixing also occurs in the nominal domain, but there is no default nominalizing affix in such cases. (5) provides examples of Greek-German (Alexiadou 2011, 2017, Alexiadou et al. 2015) and (6) illustrates Greek-English (Gardner-Chloros 2009: 50).

(5)	<i>Mixing</i>	<i>German/English</i>	<i>Greek</i>
a.	i Kél-a the.F cellar-F	der Keller the.M cellar.M	to kelar-i the.N cellar-N
b.	i Káss-a the.F cashpoint-F	die Kasse the.F cashpoint.F	to tami-o the.N cashpoint-N
c.	o Vetrét-as the.M representative.M	der Vertreter the.M representative.M	o andiprosopos the.M representative.M
d.	to matrátz-i the.N mattress.N	die Matratze the.F mattress.F	to strom-a the.N mattress-N
(6)	<i>Mixing</i>	<i>English</i>	<i>Greek</i>
a.	marketa (F)	market	agora (F)
b.	hoteli (N)	hotel	ksenodohio (N)
c.	kuka (F)	cooker	furnos (M)
d.	haspas (M)	husband	andras (M)

Given that English does not have gender marking on nouns and that the Greek exponents for gender often do not match either the gender for Greek or German, it is not plausible to argue that gender is part of the lexical specification for *Kél*. Rather, this minimal unit is best considered an abstract root that receives gender as part of being categorized as a noun (Kramer 2015, Alexiadou 2017). That is, gender can be modelled as a feature on the categorizer *n*, as illustrated in (7).



This makes gender assignment a flexible and dynamic process.

Another example that supports this line of argumentation comes from the heritage language American Norwegian, a variety of Norwegian spoken by highly dominant English speakers in the US (see Haugen 1953). (8)-(10) show the same unit surfacing both as verbs (a) and nouns (b): The unit is from English and the inflectional morphology is from Norwegian (data from the Corpus of American Nordic Speech (CANS; Johannessen 2015) and Riksem et al. 2019).

(8)	a.	fenc-a fence-PTCP 'fenced'	<i>verb</i>
	b.	en fence a.M fence 'a fence'	<i>noun</i>
(9)	a.	mow-er mow-PRS 'mows'	<i>verb</i>
	b.	mow-er-e mower-NMLZ-INDEF.PL.M 'mowers'	<i>noun</i>
(10)	a.	vota vote-PTCP 'vote'	<i>verb</i>
	b.	vot-ing-a vote-NMLZ-DEF.SG.F 'the voting'	<i>noun</i>

Given that the English units do not have the required inflectional information, we argue that these are best considered abstract roots without any grammatical information, including category. Instead, they acquire category and grammatical features through the grammatical environment in which they occur. If a root occurs together with a verbalizer, it becomes a verb, and if it occurs with a nominalizer, it becomes

a noun. Language mixing provides substantial evidence for the view that the categorizers (or other functional heads) carry all the grammatical features, the roots themselves are devoid of all grammatical information. This explains how e.g., roots from English can acquire non-native grammatical features (such as gender, or the famous double definiteness property) when they appear with Norwegian functional structure.

#### 4. Case Study 2: Language mixing in compounds

Previous research has shown that it is possible to mix words in compounds. For instance, Treffers-Daller (2005) considers mixed compounds in Brussel Dutch and how Dutch and French are mixed in this variety. She identifies three types (p. 496), as illustrated in (11).

- |         |                                       |                              |
|---------|---------------------------------------|------------------------------|
| (11) a. | velo+winkel<br>'bicycle shop'         | French non-head, Dutch head  |
| b.      | winter+paletot<br>'winter coat'       | Dutch non-head, French head  |
| c.      | gazette+marchand<br>'newspaper agent' | French non-head, French head |

In all these examples, the word order conforms to the rules of Dutch grammar: All compounds are right-headed, unlike in French. Cases of linking elements can also be found, as shown in (12).

- (12) lain+e+matrassen  
'woolen mattresses'

Importantly, French and Dutch have different structures for compounds. For this reason, Treffers-Daller argues that French elements are embedded into a Dutch compound structure. Both Dutch and French involve compounds with non-heads of the same granularity, namely phrases, i.e., roots plus functional material (see e.g., Villoing 2012 for French, Banga et al. 2013 for Dutch), thus they are not so informative regarding the role roots play in this process.

The previous literature does not really go beyond the word-level, suggesting that we need to look at different data in order to determine the actual building blocks that make up compounds. In what follows, we will look at language mixing in compounds that involve Greek, as the language has been argued to involve non-heads that are roots (Ralli 2013a).

In Greek-English, speakers produce mixed compounds as in (13) (see Seaman 1972, Alvanoudi 2019, Alexiadou 2020). English and Greek differ with respect to the nature of non-heads: non-heads in Greek must be bare roots, captured in Ralli's (2013a) *bare root constraint*, unlike in English where non-heads are phrasal (Wiese 1996, Iordăchioaia et al. 2017).

- (13) gaz-**o**-stóf-a  
'gas-LE-stove' (Seaman 1972: 196-199)

Note that the compound contains the Greek linking element, *-o-*, in addition to Greek nominal inflection on the head of the compound. As such, this example shows that speakers decompose the compound word and use the non-head root in the mixed compound.

Another option is to make use of derivational processed whereby the non-head root yields a derived word which combines with *n*. This is illustrated in (14).

- (14) grosar-ía  
'grocery store' (Seeman 1972: 197)

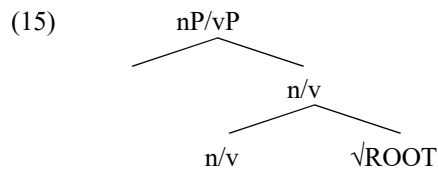
In (14), the English compound is rendered as a derived word in the mixing variety, where the non-head yields an agentive nominalization. This example illustrates that speakers can break compounds down and use a derivational process instead to conform with the relevant language requirement. This process is also known from the acquisition of compounds. Berman (2009: 313) shows that children do this during acquisition, e.g., Hebrew children say *aglan* 'wagoner' (from *agala* 'wagon') rather than *mašxan* 'puller'

for *wagon puller*. Such patterns support two claims. Firstly, they provide evidence for the view in Kroll & Stewart (1994), according to which languages share underlying concepts. The concept is lexicalized via a compound in English, but with a derived word in Greek/Hebrew. Secondly, they also support the view that derivation and compounding are part of the same grammatical domain, namely syntax. The latter view is fully in line with the main tenet of Distributed Morphology where syntactic principles extend all the way down to morphemes.

## 5. Synthesis

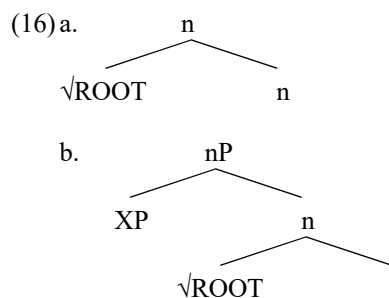
The data reviewed in this paper provide arguments in favor of the following claims: (i) Nouns and verbs are derived, they do not exist in the lexicon as primitives. Rather, the lexicon consists of a-categorial roots. (ii) Nouns and verbs emerge when a-categorial roots combine with categorizing heads (e.g., Marantz 1997, Arad 2005, Embick 2010, 2015). The morphophonological realization of these categorizers vary within and across languages. (iii) Features associated with nouns and verbs are not part of the lexical information associated with roots. This information is part of the syntactic heads that are merged above the root. More generally, multilingual speakers (and L1 acquirers) decompose words into roots and functional material and re-categorize them/use them in novel ways in the abstract structures available to them.

Let us now consider the structural correlates of these claims. We will begin with the structure of mixed nouns/verbs. The general structure looks like in (15).



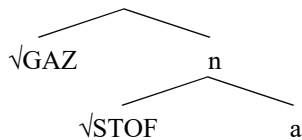
An uncategorized root merges with a categorizer, be it n or v (setting aside potential other categorizers; see Borer 2013 and Lohndal 2020 for discussion). The structural relationship between the categorizer and roots is a matter of discussion in the literature; see Alexiadou & Lohndal (2017) for a summary of the claims and arguments in favor of adjunction. In principle, the root can be from either language and so can the categorizer, although some language pairs have restrictions due to language-specific requirements (see Muysken 2000, Alexiadou 2017, and Alexiadou & Lohndal 2018 for more on this).

With respect to compounds, these have been argued elsewhere to come in at least two types: either a structure where roots act as non-heads, (16a; Harley 2009), or where phrases act as non-heads, (16b; Iordăchioaia et al. 2017). Note that for Greek, we follow Ralli (2013b) in assuming that the linker is inserted at morphophonological structure, which is why it is not included in this structure.

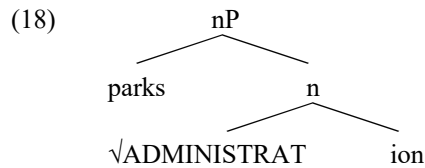


In (16a), the categorizer acts as the head. This is the structure for mixed compounds such as (13), which we illustrate in (17) (cf. Alexiadou 2020: 10).





In this structure, the mixed compound seems to have an English-based root as a non-head, which combines with a categorized root with the declension class marker *a*. It is important that the English root becomes a Greek word in order to take part in compounding. This entails that only bare roots can appear as non-heads; inflectional and derivational material cannot appear on the non-head. However, as we have seen in example (14), another strategy is also available, where the English root merges with a derivational affix before any additional inflectional information. Bilingual speakers can also use the structure in (16b), which would yield non-mixed compounds. An example is provided in (18).



That is, Greek speakers can also mix fully English compounds into their grammar.

## 6. Conclusion

Data from language contact, in particular mixing, provides crucial evidence regarding the nature of roots and the derivational mechanics of roots and functional structure. In this paper, we have argued that data from language mixing argues in favor of the following view of roots: i) roots are devoid of categorial information, ii) roots merge with (overt or covert) categorizers in the syntax, iii) the realization of roots and their functional vocabulary as identified in work on mixing informs theories of what a word is, how it is built, and how concepts are lexicalized across languages.

## References

- Áfarli, Tor A. 2015. A syntactic model for the analysis of language mixing phenomena: American Norwegian and beyond. In *Moribund Germanic Heritage Languages in North America*, B. Richard Page & Michael T. Putnam (eds.), 12-33. Leiden: Brill.
- Alexiadou, Artemis. 2011. Remarks on the morpho-syntax of code switching. In *Proceedings of the 9th International Conference on Greek Linguistics, University of Chicago, 29-31 October 2009*, Katerina Chatzopoulou, Alexandra Ioannidou & Suwon Yoon (ed.), 44-55. Columbus, OH: Ohio State University.
- Alexiadou, Artemis. 2017. Building verbs in language mixing varieties. *Zeitschrift für Sprachwissenschaft* 36: 165-192.
- Alexiadou, Artemis. 2020. Compound Formation in Language Mixing. *Frontiers in Psychology* 11:1021. doi:10.3389/fpsyg.2020.01021
- Alexiadou, Artemis, Hagit Borer & Florian Schäfer (eds.). 2014. *The Syntax of Roots and the Roots of Syntax*. Oxford: Oxford University Press.
- Alexiadou, Artemis & Terje Lohndal. 2017. The structural configurations of root categorization. In *Labels and Roots*, Leah Bauke & Andreas Blümel (eds.), 203-232. Berlin: Mouton de Gruyter.
- Alexiadou, Artemis & Terje Lohndal. 2018. Units of language mixing: A cross-linguistic Perspective. *Frontiers in Psychology* 9:1719. doi:10.3389/fpsyg.2018.01719
- Alexiadou, Artemis, Terje Lohndal, Tor A. Áfarli & Maren Berg Grimstad. 2015. Language mixing: a Distributed Morphology approach. In *Proceedings of the Forty-Fifth Annual Meeting of the North East Linguistic Society*, Thuy Bui & Deniz Özyildiz (eds.), 25-38. Amherst, MA: GSLA.
- Alvanoudi, Angeliki 2019. *Modern Greek in Diaspora: An Australian Perspective*. London: Palgrave.
- Arad, Maya. 2005. *Roots and Patterns: Hebrew Morpho-Syntax*. Dordrecht: Springer.
- Banga, Arina, Esther Hanssen, Anneke Neijt & Robert Schreuder. 2013. Preference for linking element *-en-* in Dutch noun-noun compounds: native speakers and second language learners of Dutch. *Morphology* 23: 33-56.
- Berman, Ruth. 2009. Children's acquisition of compound constructions. In *The Oxford Handbook of Compounding*, Rochelle Lieber & Pavol Štekauer (eds.), 298-322. New York: Oxford University Press.

- Borer, Hagit. 2013. *Structuring Sense: Taking Form*. Oxford: Oxford University Press.
- Borer, Hagit. 2014. The category of roots. In *The Syntax of Roots and the Roots of Syntax*, Artemis Alexiadou, Hagit Borer & Florian Schäfer (eds.), 111-148. Oxford: Oxford University Press.
- Callies, Marcus & Christel Stolz (eds.). 2016. Word-formation. Language contact, language contrast and language comparison. *STUF, Language Typology and Universals* 69(4).
- Embick, David. 2010. *Localism versus Globalism in Morphology and Phonology*. Cambridge, MA: MIT Press.
- Embick, David. 2015. *The Morpheme*. Berlin: Mouton de Gruyter.
- Gardner-Chloros, Penelope. 2009. *Code-Switching*. Cambridge: Cambridge University Press.
- Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, Kenneth Hale & Samuel Jay Keyser (eds.), 111-176. Cambridge, MA: MIT Press.
- González-Vilbazo, Kay & Luis López. 2011. Some properties of light verbs in code switching. *Lingua* 121: 832-850.
- Grimstad, Maren Berg, Brita Ramsevik Riksem, Terje Lohndal & Tor A. Åfarli. 2018. Lexicalist vs. exoskeletal approaches to language mixing. *The Linguistic Review* 35: 187-218.
- Harley, Heidi. 2009. Compounding in Distributed Morphology. In *The Oxford Handbook of Compounding*, Rochelle Lieber & Pavol Štekauer (eds.), 129-144. New York: Oxford University Press.
- Haugen, Einar. 1953. *The Norwegian Language in America*. Philadelphia: University of Philadelphia Press.
- Iordăchioaia, Gianina, Artemis Alexiadou & Andreas Pairamidis. 2017. Morphosyntactic sources for nominal synthetic compounds in English and Greek. *Journal of Word Formation* 1: 47-72.
- Johannessen, Janne Bondi. 2015. The Corpus of American Norwegian Speech (CANS). In *Proceedings of the 20th Nordic Conference of Computational Linguistics, NODALIDA 2015, May 11-13, 2015*, B. Megyesi (ed.), 239-243. Vilnius: Linköping University Electronic Press.
- Kramer, Ruth. 2015. *The Morphosyntax of Gender*. Oxford: Oxford University Press.
- Kroll, Judith F. & Erika Stewart. 1994. Category interference in translation and picture naming: Evidence for asymmetric connection between bilingual memory representations. *Journal of Memory and Language* 33: 149-174.
- Lohndal, Terje. 2020. Syntactic Categorization of Roots. *Oxford Research Encyclopedia in Linguistics*. Doi: 10.1093/acrefore/9780199384655.013.257
- López, Luis. 2020. *Bilingual Grammar. Toward an Integrated Model*. Cambridge: Cambridge University Press.
- MacSwan, Jeff. 2005. Codeswitching and generative grammar: a critique of the MLF model and some remarks on "modified minimalism". *Bilingualism: Language and Cognition* 8: 1-22.
- Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *University of Pennsylvania Working Papers in Linguistics* 4: 210-225.
- Muysken, Pieter. 2000. *Bilingual Speech: A Typology of Code-Mixing*. Cambridge: Cambridge University Press.
- Muysken, Pieter. 2013. Language contact outcomes as the result of bilingual optimization strategies. *Bilingualism: Language and Cognition* 16: 709-730.
- Poplack, Shana. 1980. "Sometimes I'll start a sentence in Spanish Y TERMINO EN ESPANOL": toward a typology of code-switching. *Linguistics* 18: 581-618.
- Ralli, Angela. 2013a. *Compounding in Modern Greek*. New York: Springer.
- Ralli, Angela. 2013b. Compounds and its locus of realization. *Word Structure* 6: 181-200.
- Riksem, Brita Ramsevik, Maren Berg Grimstad, Terje Lohndal & Tor A. Åfarli 2019. Language mixing within verbs and nouns in American Norwegian. *Journal of Comparative Germanic Linguistics* 22: 189-209.
- Sankoff, David & Shana Poplack. 1981. A formal grammar for code-switching. *Paper in Linguistics* 14: 3-45.
- Seaman, P. David. 1972. *Modern Greek and American English in Contact*. The Hague: Mouton.
- Treffers-Daller, Jeanine. 2005. Evidence for insertional codemixing: mixed compounds and French nominal groups in Brussels Dutch. *International Journal of Bilingualism* 9: 477-506.
- Villoing, Florence. 2012. French Compounds. *Probus* 24: 29-60.
- Wiese, Richard. 1996. Phrasal compounds and the theory of word syntax. *Linguistic Inquiry* 27: 183-193.
- Wohlgemuth, Jan. 2009. *A Typology of Verbal Borrowings*. Berlin: Mouton de Gruyter.